The University of Nottingham

School of Molecular Medical Sciences

Faculty of Medicine and Health Sciences

Molecular Genetic Studies in Pregnancies Affected by Preeclampsia and Intrauterine Growth Restriction

by

Ayat Abdel-Rahman Sayed Abd-Rabou

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

August 2011

Abstract

Preeclampsia and fetal growth restriction (FGR) are common and costly obstetric complications. Both conditions are associated with immediate and remote mortality and morbidity for the mother and the offspring. Impaired placentation and aberrant maternal systemic responses are implicated as pathophysiological mechanisms in preeclampsia and FGR. Both preeclampsia and FGR are known to have a clear genetic basis. This study has investigated the roles of several candidate genes including those previously associated with diabetes (TCF7L2, FTO, PPAR-g, CDKN2B-AS1 and KCNJ11), and epidermal growth factor (EGF). Functional consequences of variants within the EGF gene were also investigated.

A bidirectional association between type 2 diabetes (T2D) and preeclampsia is consistently reported, whereby each condition is associated with an increased risk of the other. Furthermore, fetal growth restriction, which complicates 30% of preeclamptic pregnancies, predisposes the offspring to an increased risk of type 2 diabetes and coronary artery disease (CAD) later in life. 11 single nucleotide polymorphisms (SNPs) reproducibly associated with T2D in the TCF7L2, FTO, PPAR-γ, CDKN2B-AS1 and KCNJ11 genes were investigated as susceptibility loci for preeclampsia and fetal growth restriction in a maternal case control study. The study group consisted of 448 white western European women with preeclampsia, 673 controls with no evidence of preeclampsia, 243 women with pregnancies complicated by FGR, and 570 controls with no evidence of growth restriction. A maternal haplotype on the T2D region of the CDKN2B-AS1 gene on chromosome 9p21 was found to be a risk variant for fetal growth

restriction (P=0.005). The other 9 investigated SNPs in TCF7L2, FTO, PPAR-γ, and KCNJ11 showed no association with growth restricted pregnancies. None of the SNPs investigated showed an association with preeclampsia. These findings suggest that some maternal diabetogenic risk variants are associated with an altered risk of FGR pregnancy but not preeclampsia. The results require replication in a larger sample and fetal-maternal gene interactions merit investigation.

Epidermal growth factor (EGF) is described as a major regulator of the placentation process. It also helps to maintain an adequate blood supply to the growing fetus through its effects on umbilical vessel tone. Investigating the role of two genetic variants of the EGF gene in susceptibility to preeclampsia and FGR showed that the maternal variants, rs4444903 in the 5'UTR and rs2237051 in exon 14 of the EGF gene has no effect on the risk of preeclampsia or FGR pregnancy. The G allele of the SNP rs4444903 was associated with higher systolic blood pressure measures in the control group.

The G allele of the rs4444903 and the A allele of rs2237051 have been associated with increased risk for FGR and lower birth weight in a previous study from our laboratory. This led to investigations to characterize the functional consequences of the two SNPs in the EGF gene on transcription, translation and ribonucleic acid (RNA) splicing using a variety of methods. These experiments have shown that the G allele of rs4444903 was transcriptionally more active than the A allele in hepatocellular carcinoma (HepG2) and more active than EGFP on its own in choriocarcinoma (Jeg-3) cell lines using a luciferase reporter gene assay. There was no effect of this variant on

translational efficiency in the cell lines investigated using reporter gene assays, or in a cell free environment using an *in vitro* translation assay. DNA-protein interaction was investigated using nuclear extract from HepG2 cells to further define the mechanism by which the G allele exerts its higher transcriptional activity. Initial experiments suggest that the Sp1 transcription factor interacts with and represses the A allele of the rs4444903 SNP. The study also demonstrated no evidence of higher activity of the G allele on EGF expression *in vivo* using term placental tissues. It was expected that higher EGF expression as a function of genotype at rs4444903 SNP may lead to down regulation of the EGFR in the placenta, which was not confirmed in this study. SNP rs2237051 in exon 14 of the EGF gene is in strong linkage disequilibrium with rs4444903, and disrupts a predicted exon splicing enhancer region. This polymorphism was investigated using a minigene assay, but there was no evidence that it affected splicing of exon 14. Taken together, these findings provide no evidence that EGF genetic variants alter the risk of preeclampsia or FGR though functioning.

Abstracts and scientific communications arising from this work

Ayat Sayed, Sally Chappell, Linda Morgan. The role of diabetes susceptibility genetic variants in preeclampsia. Pregnancy Hypertension 1, Supplement 1 (2010) S71. XVII ISSHP World Congress, Melbourne, Australia, 3-6 October 2010.

Ayat Sayed, Sally Chappell, Linda Morgan. Maternal type 2 diabetes risk alleles and intrauterine growth. The International Society of Obstetric Medicine Meeting, Melbourne, Australia, 1-3 October 2010.

Ayat Sayed, Sally Chappell, Linda Morgan. The role of epidermal growth factor (EGF) Polymorphism in preeclampsia. J Med Genet 2010; 47: Supplement 1, S84. British Human Genetics Conference, University of Warwick, 6-8 September 2010.

Ayat Sayed, Sally Chappell, Linda Morgan. Functional characterization of the EGF 5'UTR genetic variation. International Society for the Study of Hypertension in Pregnancy. Oxford, UK, 11-12 September 2009.

Ayat Sayed, Sally Chappell, Linda Morgan. Using A Minigene Assay to Assess The Function of A Polymorphism Associated with Intra-uterine Growth Restriction. The Anatomical Society of Great Britain & Ireland Summer Meeting, the University of Nottingham, 2-4 July, 2008.

Hanna Peterson, Katja Kivinen, Leena Hiltunen, Elina Salmela, Tuuli Lappalainen, Vesa Rasi, Ayat Sayed, Lucilla Poston , Shaun P Brennecke, Matthew P Johnson, Linda Morgan, Eric K Moses, Juha Kere and Hannele Laivuori. Common variants on chromosome 9p21 are associated with pre-eclampsia in the Finnish population. (Submitted for publication).

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors, Dr. Linda Morgan and Dr. Sally Chappell for all their guidance and support throughout this demanding research process. Without their commitment and enthusiasm, this work will not be completed. I only hope my efforts have matched their input.

I would also like to thank Dr Tamar Guetta-Baranes, who has given me immense support throughout my study. Thanks to Dr Paula Williams, who helped me in immunostaining of the placental tissues, and thanks to Dr Hiten Mistry, who collected the placental samples, Dr Kristell Brown for her help and support. A big thank you goes to Haitham Qutob, and Aurore van Hoof who helped me in the genotyping. Thanks to Dr Linda Morgan, and Dr Clare Tower for the collection of DNA samples they provided. Thanks for the mothers who participated in the study, and for everyone who helped in their recruitment.

I acknowledge the general help and support from all the members of Human Genetics laboratory, and members of the School of Molecular Medical Sciences both past and present, for all their technical and much needed moral support. I would also like to thank friends outside the department especially PhD students in the Mol postgraduate room, my Egyptian friends in Nottingham, specially Tarek and Maram.

I would also like to express my sincere gratitude to the Medical Biochemistry Department, Assiut University, and the Egyptian Ministry of Higher Education for granting my PhD project and also thanks to the Egyptian cultural and educational bureau in London who successfully substituted my official work during the grant periods of the work in UK.

I am forever indebted to my Mum and Dad for understanding, endless patience and encouragement when it was most required. A special compliment goes to my husband, Hatem, and my sons, Muhammad and Mustafa, who are keeping me alive. Hatem sacrificed the most to make this whole work possible and it is his name really deserves to be along side mine on the front. Finally, this thesis is dedicated to the memory of my mother-in-law, whose love and care I missed.

Abbreviations

AA	Amino acids
AGA	Appropriate for gestational age
AGT	Angiotensin
AGTR2	Angiotensin II Type 2 Receptor
AGTR1	Angiotensin II Type 1 Receptor
Ala	Alanine
BMI	Body mass index
Bp	Base pair
BWC	Birth weight centile
CAD	Coronary artery disease
CDKN2A	Cyclin dependent kinase inhibitor 2A
C-TBP-1	C terminal binding protein -1
CI	Confidence interval
CRP	C reactive protein
CT	Cytotrophoblast
DBD	DNA-binding domain
15d-PGJ2	15-deoxy-12, 14-prostaglandin J_2
DNA	Deoxyribonucleic Acid
ddNTP	Dedeoxynucleoside triphosphate
dNTP	Deoxynucleoside triphosphate
Dvl	Dishevelled
EC	Endothelial cell
EDTA	Ethylenediaminetetraacetic acid
EFW	Estimated fetal weight
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMSA	Electrophoretic Mobility Shift Assay
ESE	Exonic splicing enhancers

ESS	Exonic splicing silencer
EVT	Extra-villous trophoblast
FasL	Fas ligand
FGR	Fetal growth restriction
FTO	Fat mass- and obesity-associated gene
GCK	Glucokinase gene
GDM	Gestational diabetes mellitus
GLP-1	Glucagon like peptide -1
GSK-3	Glycogen synthase kinase 3
GWAS	Genome wide association studies
hCG	Human chorionic gonadotrophin
HDL	high density lipoprotein
HELLP	Haemolysis, Elevated Liver enzymes and Low Platelet count
HIF-1	Hypoxia inducible factor-1
HOMA-IR	Homeostasis Model of Assessment-Insulin Resistance
hPL	Human placental lactogen
hnRNP	Heterogeneous nuclear ribonucleoprotein protein
IGF-I	Insulin-like growth factor I
IGF-II	Insulin-like growth factor II
IR	Insulin resistance
IVT	In vitro transcription and translation
Kb	Kilobase
kDa	Kilo Dalton
KCNJ11	Potassium inwardly-rectifying channel, subfamily J, member 11
LD	Linkage disequilibrium
LBD	Ligand-binding domain
LDL	Low density lipoprotein
LGA	Large for gestational age
MAF	Minor allele frequency

Mb	Megabase
MMPs	Matrix metalloproteinase (s)
mRNA	Messenger ribonucleic acid
MS-PCR	Mutagenically separated PCR
MTHFR	Methylene tetrahydrofolate reductase
NO	Nitric oxide
NOS3	Endothelial nitric oxide synthase
OGTT	Oral glucose tolerance test
OR	Odds ratios
ORF	Open reading frame
PAR	Population attributable risk
PGE	Prostaglandin E
PGF	Placental growth factor
PPAR-γ	Peroxisome proliferator-activated receptor gamma
PPREs	Peroxisome proliferator-activated receptor response elements
Pro	Proline
PPT	Precipitation
r^2	Linkage correlation coefficient
ROS	Reactive oxygen species
RT-PCR	Reverse transcriptase-polymeras chain reaction
RXR	Retinoic X receptors
sEng	Soluble endoglin
sFlt-1	Soluble fms-like tyrosine kinase-1
SHBG	Sex hormone binding globulin
SNP	Single nucleotide polymorphism
snRNP	Small nuclear ribonucleoproteins
Sp1	Specificity protein 1
SR proteins	Serine–arginine rich proteins
SRE	Splicing regulatory element

SS	Splice site
ST	Syncytiotrophoblast
STMPs	syncytio-trophoblast microparticles
T2D	Type 2 diabetes
TAE	Tris-acetate EDTA
TBE	Tris-borate EDTA
TCF7L2	Transcription factor 7 - like 2
TDT	Transmission-disequilibrium testing
TGF-α	Transforming growth factor-α
TGF-ß	Transforming growth factor ß
TNF-α	Tumour necrosis factor-α
TRAIL	TNF-related apoptosis inducing ligand
TSS	transcription start sites
uORF	upstream open reading frames
5'UTR	5' untranslated regions
VEGF	Vascular endothelial growth factor
VIP	Vitamins in Preeclampsia
WNT	Wingless pathway
WTCCC	Wellcome Trust Case Control Consortium

Table of contents

ABSTRACT 2

ABSTRACTS AND SCIENTIFIC COMMUNICATIONS ARISING FROM THIS WORK 5

ACKNOWLEDGEMENTS 6

ABBREVIATIONS 7

TABLE OF CONTENTS 11

LIST OF FIGURES 15

LIST OF TABLES 16

PREFACE 18

1 INTRODUCTION 20

1.1 Preeclampsia 21
1.1.1 Epidemiology of preeclampsia 21
1.1.2 Clinical manifestations of preeclampsia 26
1.1.3 Risk factors for preeclampsia 30
1.1.4 The role of the placenta 33
1.1.4.1 Insufficient spiral artery remodelling 34
1.1.4.2 Abnormalities of the villous trophoblast 37
1.1.5 Maternal systemic changes during pregnancy and preeclampsia 41
1.1.6 The relation between abnormal placentation and systemic maternal manifestations 45
1.1.7 Genes and preeclampsia 46
1.1.7.1 Heritability of preeclampsia 46
1.1.7.2 Genetic studies in preeclampsia 48
1.1.7.3 Genetic linkage studies 49
1.1.7.4 Candidate gene studies 51

1.2 Fetal growth restriction 60
1.2.1 Definition of FGR 61
1.2.2 Implication of FGR 64
1.2.3 Aetiology and risk factors for FGR 66
1.2.4 The role of the placenta in FGR 69
1.2.5 Genetics of FGR 71
1.2.6 Genetic studies of FGR 72
1.2.6.1 Fetal programming and diabetes genes 73
1.2.6.2 Genomic imprinting defects and FGR 77
1.2.6.3 Other candidate gene studies 78

1.3 Summary and hypothesis 79

2 TYPE 2 DIABETES SUSCEPTIBILITY GENES AND THE RISK OF PREECLAMPSIA AND FETAL GROWTH RESTRICTION 81

2.1 Introduction 81

2.2 Candidate genes 83
2.2.1 Transcription factor 7 - like 2 (TCF7L2) 83
2.2.2 Introduction 83
2.2.2.1 Genomic structure 84
2.2.2.2 Polymorphisms of the TCF7L2 86
2.2.2.3 Role of TCF7L2 gene in diabetes 89
2.2.2.4 Role of TCF7L2 variants in pregnancy disorders 90
2.2.3 Fat Mass- And Obesity-Associated Gene (FTO) 92
2.2.3.1 Introduction 92
2.2.3.2 Genomic structure and polymorphisms of FTO gene 92
2.2.3.3 Role of FTO in Diabetes, obesity and metabolic syndrome 93
2.2.3.4 FTO variants and CAD risk 96
2.2.3.5 Functional effects of the FTO variants 97
2.2.3.6 Role of FTO in pregnancy 98
2.2.4 Peroxisome proliferator-activated receptors-gamma (PPAR-γ) 99
2.2.4.1 Introduction 99
2.2.4.2 Genomic structure of PPAR-γ 102
2.2.4.3 Polymorphisms of the PPAR-γ 105
2.2.4.4 Mechanism of action of PPAR-γ variants in T2D 109
2.2.4.5 The role of PPAR-γ in pregnancy 110
2.2.5 Cyclin dependent kinase inhibitor 2 B- (CDKN2B) antisense RNA 1 (CDKN2B-AS1) 114
2.2.5.1 Introduction 114
2.2.5.2 Genomic structure of CDKN2B-AS1 115
2.2.5.3 Polymorphisms of CDKN2B-AS1 116
2.2.5.4 Functional effects of the CAD risk variant of CDKN2B-AS1 121
2.2.6 Potassium Inwardly-Rectifying Channel, Subfamily J, Member 11, KCNJ11 122
2.2.6.1 Introduction 122
2.2.6.2 Genomic structure and polymorphism of the KCNJ11 124
2.2.6.3 Role of KCNJ11 variants in type 2 diabetes 126
2.2.6.4 Role of KCNJ11 variants in pregnancy disorders 129

2.3 Aim of the study 130

2.4 Methods 131
2.4.1 Patient recruitment 131
2.4.1.1 Vitamins in Preeclampsia (VIP) trial 132
2.4.1.2 Nottingham FGR collection 133
2.4.1.3 Nottingham preeclampsia collection 134
2.4.2 Preeclampsia and fetal growth restriction phenotype definition 136
2.4.3 Power calculation 137
2.4.4 DNA extraction 137
2.4.5 Genotyping 138
2.4.5.1 The Kbioscience Competitive Allele Specific PCR (KASPr) 139
2.4.5.2 Mutagenically separated PCR (MS-PCR) 142
2.4.5.3 PCR reaction and condition 146
2.4.5.4 Gel electrophoresis 147
2.4.5.5 DNA sequencing 148
2.4.6 Genotyping quality control 149
Several procedures were put in place to ensure that the genotyping was carried out correctly. 149

2.4.6.1 Quality control for Kbioscience genotyping 149
2.4.6.2 Quality control for in house genotyping 149
2.4.7 Statistical analysis 150

2.5 Results 151
2.5.1 Quality control and population genetics results 151
2.5.2 Preeclampsia cases and controls: Phenotypic Data 153
2.5.3 Preeclampsia case control genetic association study 157
2.5.4 Fetal growth restriction (FGR) phenotypic data 161
2.5.5 Fetal growth restriction (FGR) case control genetic association study 165
2.5.6 The effect of T2D variants on maternal BMI and fetal birth weight 169

2.6 Discussion 172
2.6.1 Type 2 diabetes genes in preeclampsia case control study 172
2.6.2 Type 2 diabetes genes in FGR case control study 179

3 ROLE OF EPIDERMAL GROWTH FACTOR GENETIC VARIANTS IN PREECLAMPSIA AND FETAL GROWTH RESTRICTION 184

3.1 Introduction 184

3.2 Aims of the study 186

3.3 Genetic association study 187
3.3.1 Introduction 187
3.3.1.1 Epidermal growth factor and placental development 187
3.3.1.2 EGF and fetal growth 189
3.3.1.3 Epidermal growth factor levels during pregnancy 191
3.3.1.4 Polymorphisms of the EGF gene and the risk of preeclampsia and FGR 192
3.3.2 Methods 193
3.3.3 Results 194
3.3.3.1 Association of EGF variants with preeclampsia 195
3.3.3.2 Association of EGF variants with FGR 195
3.3.3.3 Association of EGF variants with maternal blood pressure 198
3.3.3.4 Association of EGF variants with offspring birth weight 199
3.3.4 Discussion 201

3.4 The effect of c.2556G>A (rS2237051) on EGF mRNA splicing 205
3.4.1 Introduction 205
3.4.1.1 Role of mRNA splicing in human diseases 207
3.4.1.2 Analysis of splicing regulatory elements 208
3.4.2 Methods 210
3.4.2.1 EGF exon 14 sequence bioinformatic analysis 210
3.4.2.2 Minigene constructs 210
3.4.2.3 Cell culture and transfection 212
3.4.2.4 RNA extraction and RT-PCR 213
3.4.3 Results 215
3.4.3.1 Initial in silico sequence analysis 215
3.4.3.2 In vivo splicing assay 218
3.4.4 Discussion 221

3.5 Functional Characterization of the EGF 5'UTR genetic variation: translational versus transcriptional effects 224

3.5.1 Introduction 224
3.5.2 Hypothesis and aims 226
3.5.2.1 Luciferase reporter gene assay 227
3.5.2.2 In vitro transcription and translation assay (IVT) 231
3.5.2.3 Electrophoretic Mobility Shift Assay (EMSA) 232
3.5.3 Methods 234
3.5.3.1 Cloning of the 5'UTR of the EGF-luciferase reporter constructs 234
3.5.3.2 Cloning of the EGF promoter and EGF promoter/5'UTR fragments 237
3.5.3.3 Cell culture and transfection 241
3.5.3.4 In vitro transcription and translation 244
3.5.3.5 Electrophoretic mobility shift assay, EMSA 247
3.5.3.6 Bioinformatics 249
3.5.3.7 Data analysis 250
3.5.4 Results 251
3.5.4.1 Optimization of Jeg-3 cells transfection conditions 251
3.5.4.2 Effect of EGF 5'UTR on reporter gene expression 253
3.5.4.3 In vitro transcription and translation 261
3.5.4.4 Electrophoretic mobility shift assay 266
3.5.5 Discussion 270

3.6 The effect of SNP rs4444903 on EGF and EGFR expression in the placenta 274
3.6.1 Introduction 274
3.6.2 Signal transduction by EGFR 274
3.6.3 Internalization and down regulation of the EGFR 276
3.6.4 EGF-EGFR in trophoblast (in normal and abnormal pregnancy) 278
3.6.5 Aim of the study 281
3.6.6 Methods 282
3.6.6.1 Sample collection 282
3.6.6.2 Immunohistochemical staining for EGFR and EGF 282
3.6.6.3 Analysis of the EGF and EGFR immunostaining of the placenta 284
3.6.6.4 Genotyping of the placental samples 284
3.6.6.5 Statistical analysis 285
3.6.7 Results 286
3.6.8 Discussion 292

3.7 Conclusion 295

4 FINAL DISCUSSION AND FUTURE WORK 296

4.1 Summary 296

4.2 Strengths and limitations of the study 299

4.3 Further work 302

5 REFERENCES 306

List of Figures

Figure 1-1: Preeclampsia as a life long risk for cardio-metabolic disorders 26
Figure 1-2: Spiral artery remodelling in normal and hypertensive pregnancy 37
Figure 1-3: Some of the candidate gene studied in preeclampsia 58
Figure 1-4: The effect of maternal and paternal diabetes genes on birth weight 75
Figure 2-1: Canonical B-catenin pathway 84
Figure 2-2: Genomic structure of the TCF7L2 85
Figure 2-3: Organization of the FTO gene 93
Figure 2-4: Genomic structure of PPAR-γ 104
Figure 2-5: Schematic presentation of chromosome 9p21 region 116
Figure 2-6 : The pattern of LD in chromosome 9p21 region 120
Figure 2-7: Structure of KATP channel in the b-cell of the pancreas 124
Figure 2-8: LD blocks and SNPs in both KCNJ11 (Kir6.2) and ABBC8 (SUR1) genes 125
Figure 2-9: Number and frequency of participant women from different collections 131
Figure 2-10: KASPar genotyping technology 140
Figure 2-11: SNPviewer of samples genotyped for SNP rs1333049 141
Figure 2-12: MS-PCR, three-primer method 144
Figure 2-13: MS-PCR, four-primer method 145
Figure 2-14: Electophoresis banding pattern for SNP rs7903146 151
Figure 2-15: Electrophoresis banding pattern for SNP rs1801282 152
Figure 2-16: Electropherogram showing CC and TT genotypes for SNP rs7903146 152
Figure 3-1: Genomic structure of the epidermal growth factor gene 185
Figure 3-2: Core and cis- regulatory splicing elements 206
Figure 3-3: Outline of the minigene technique 209
Figure 3-4: ESE-Finder graphical output for exon 14 217
Figure 3-5: Minigene constructs and *in vivo* splicing assay 219
Figure 3-6: Features by which 5'UTRs may reduce the efficiency of translation 225
Figure 3-7: Firefly and Renilla luciferase mediated reactions 227
Figure 3-8: Dual luciferase assay 229
Figure 3-9: pGL3 Control vector 230
Figure 3-10: pcDNA3.1 (+) vector 232
Figure 3-11: Plasmid constructs used in the dual luciferase assay and in vitro transcription 236
Figure 3-12: BP reaction 239
Figure 3-13: LR reaction 240
Figure 3-14: Optimization of the Jeg-3 cells transfection 252
Figure 3-15: Relative luciferase activity for the EGF 5'UTR A and G alleles in pGL3C and pGL3P 254
Figure 3-16: Relative luciferase activity from EGF 5'UTR A and G alleles in pGL3B constructs 255
Figure 3-17: Fold change in the activity of G allele/A allele of the 5'UTR of the EGF gene in different pGL3 constructs 256
Figure 3-18: Fold change in the activity of EGF 5'UTR A and G alleles with and without the EGF promoter relative to the EGF promoter alone 259
Figure 3-19: In vitro transcription reaction on 1% agarose gel 261
Figure 3-20: Luciferase output from A and G alleles of the 5'UTR of the EGF gene after in vitro transcription and translation 262
Figure 3-21: Luciferase output showing the effect of EGF 5'UTR A and G alleles on the translation of the downstream gene 263

Figure 3-22: mRNA secondary structure of the first 120bp of the EGF 5'UTR. 264
Figure 3-23: mRNA secondary structure of the full length EGF-5'UTR 265
Figure 3-24: Prediction of the transcription factors binding to 5'UTR of the EGF. 266
Figure 3-25: Electrophoretic mobility shift assays for 5'UTR of the EGF gene, and Sp1 site ...268
Figure 3-26: Supershift assays using antibodies to Sp1 with the 5'UTR of the EGF gene 269
Figure 3-27: The EGFR-signalling pathway 276
Figure 3-28: Ligand-regulated internalization and degradation of EGFR 278
Figure 3-29: EGFR staining of the placenta 289
Figure 3-30: EGF staining of the placenta 290
Figure 3-31: EGF and EGFR immunostaining in placentas from normal and preeclampsia-FGR pregnancies 291
Figure 3-32: EGF and EGFR immunostaining in placentae according to genotype at marker rs4444903 of the EGF gene 291

List of tables

Table 1-1: Classification and Diagnostic Features of Hypertension in Pregnancy 30
Table 1-2: Fetal and maternal risk factors for FGR pregnancy 67
Table 2-1: T2D genes included in the present study 82
Table 2-2: Clinical and demographic data for participants in preeclampsia and FGR genetic case control studies stratified by collection 135
Table 2-3: Power calculations for case control studies 137
Table 2-4: Primers used for MS-PCR 143
Table 2-5: MS-PCR reaction for SNP rs7903146, and rs18012182 146
Table 2-6: PCR and electrophoresis conditions for SNP rs7903146 and rs1801282 147
Table 2-7: Minor allele frequency and quality control data for study variants 153
Table 2-8: Phenotypic data for preeclampsia and healthy pregnancy 154
Table 2-9: Parity, smoking habit and baby's sex in preeclampsia and healthy pregnancy 155
Table 2-10: Frequency of risk factors in preeclampsia and control groups recruited from the VIP collection 156
Table 2-11: Rate of preeclampsia in women with risk factors recruited from the VIP collection 157
Table 2-12: Genotype and allele distributions of T2D alleles and corresponding odds ratios for preeclampsia 158
Table 2-13: Haplotype frequency for T2D susceptibility genes in preeclampsia cases and controls 160
Table 2-14: Phenotypic data for FGR and healthy pregnancy 162
Table 2-15: Parity, smoking habit and baby's sex in FGR and healthy pregnancy 162
Table 2-16: Frequency of risk factors FGR and control groups recruited from VIP collection ..163
Table 2-17: Rate of FGR in women with risk factors within the VIP collection 164
Table 2-18: Genotype and allele distributions of T2D alleles and corresponding odds ratios for fetal growth restriction (FGR) 166
Table 2-19: Haplotype frequency of type 2 diabetes genes in FGR cases and control 168
Table 2-20: Effect of genetic variants on maternal body mass index and baby's birth weight in control groups 170
Table 3-1: Genotyping quality control indices 194

Table 3-2: Genotype and allele distributions for EGF variants and corresponding odds ratios for preeclampsia 196
Table 3-3: Genotype and allele distributions of the EGF variants and corresponding odds ratios for fetal growth restriction (FGR) 197
Table 3-4: Haplotype frequency of the EGF variants in preeclampsia and fetal growth restriction cases and controls. 198
Table 3-5: Maternal systolic and diastolic blood pressure and fetal birth weight as a function of the genotype of EGF variants in the control group. 200
Table 3-6: Tabular presentation showing the effect of rs2237051 G>A on different cis acting regulatory elements in the vicinity of the polymorphism 220
Table 3-7: Primers used for experiments in 3.5.1 237
Table 3-8: Electrophoretic mobility shift assay for 5'UTRA, 5'UTRG of the EGF gene and Sp1 probes 249
Table 3-9: Relative luciferase activity for A and G alleles of the EGF 5'UTR in different pGL3 vectors in HepG2 and Jeg-3 cells. 257
Table 3-10: The activity of EGF 5'UTR A and G alleles with and without the EGF promoter relative to the EGF promoter alone 260
Table 3-11: Details of antibodies and positive control tissues used in this study 283
Table 3-12: Demographic and pregnancy characteristics of subject groups 286

Preface

This thesis describes some genetic and molecular genetic aspects of preeclampsia and fetal growth restriction in a UK population.

The introductory chapter is presented in 2 sections. The first is related to preeclampsia and the second to fetal growth restriction. In each of these sections, the magnitude of the disease and its impact on public health and the development of future cardio-metabolic disorders including type 2 diabetes (T2D), and coronary artery disease (CAD) is highlighted. Definitions and diagnosis, risk factors, and pathological features involving the placenta and the maternal systemic disturbances are also described. Special attention is drawn to the genetics of preeclampsia and fetal growth restriction, presenting evidence for a genetic basis and the approaches and the future directions genetic investigations may take. This chapter concludes with a statement of the hypothesis and the objectives of the study.

The second chapter of this thesis describes the genetic case control association study, investigating diabetes susceptibility genes and their role in preeclampsia and fetal growth restriction. This chapter comes in 4 main sections. The first section gives an overview of the 5 candidate genes to be investigated (TCF7L2, FTO, PPARG, CDKN2B-AS1 and KCNJ11. The second section of chapter 2 describes patient recruitment, genotyping and data analysis. The third and fourth sections present the results and discussion of this investigation.

The third chapter is devoted to epidermal growth factor (EGF). This chapter comprises 4 sections each with its own introduction, methods, results and discussion. The first section describes the EGF case control association study. The second and third sections describe the functional impact of rs2237051 and rs4444903 SNPs on mRNA splicing, transcription and translation of the EGF gene respectively. The fourth section shows how the expression of the EGF and EGFR in the placental tissues may vary as a function of genotype at rs4444903 SNP.

Chapter 4 provides a summary of all the experimental work, a discussion of the strengths and limitations of the study and the prospects for future work.

1 Introduction

Preeclampsia and fetal growth restriction (FGR) are two related pregnancy disorders, in which fetal growth restriction complicates up to 30% of preeclampsia cases. Both conditions carry substantial maternal and perinatal morbidity and mortality. Preeclampsia and FGR are characterized by familial aggregation and high recurrence rates among relatives, with a clear genetic basis. There is substantial evidence that both maternal and fetal genes are involved in the risk for preeclampsia and FGR.

Reduced uteroplacental blood flow due to abnormalities in vascular remodelling of the maternal blood vessels and shallow trophoblast invasion are the pathological features that characterize the placentas from preeclampsia and fetal growth restriction. Signs of placental ischaemia are also reported in isolated cases of fetal growth restriction. However, preeclampsia is a syndrome that is not confined to the placenta, but involves all the maternal organs and systems, and a proportion of cases of preeclampsia are free from placental abnormalities. This indicates that although related, both conditions still have distinct features. This introductory chapter will deal with preeclampsia and fetal growth restriction separately, highlighting their impact on maternal and fetal health, the diagnosis, the pathology, and the genetics of both conditions, and will highlight the shared and disparate feature between preeclampsia and FGR.

1.1 Preeclampsia

Preeclampsia is a syndrome affecting 3-5% of pregnant women worldwide, manifested by new onset hypertension and proteinuria at or after 20 weeks of gestation (Redman and Sargent, 2005). In its most severe forms it is a life threatening condition, and carries both immediate and long term adverse consequences to the mother.

1.1.1 Epidemiology of preeclampsia

Preeclampsia is a leading cause (12%) of maternal death worldwide, and contributes to around 16% of maternal deaths in the developed countries. In developing countries, where 99% (35500) of global maternal deaths occur, preeclampsia comes after birth related haemorrhage and sepsis as a cause of maternal death (Khan et al., 2006, Duley, 2009). Preeclampsia is also associated with fetal and neonatal complications, including FGR in 30% of cases, and an increased rate of prematurity as a result of expedited delivery as a treatment for severe preeclampsia. Other neonatal complications include perinatal death, and long-term cardiovascular morbidity associated with low birth weight (Sibai et al., 2005).

The spectrum of the problem extends beyond the index pregnancy. Women affected by preeclampsia are at increased risk of developing preeclampsia in subsequent pregnancies, which may influence maternal decisions about future pregnancies. The recurrence risk for preeclampsia in a subsequent pregnancy ranges from 7.5% to 65%, depending on sample size and inclusion criteria. Trogstad et al (2004) reported that the risk of preeclampsia in a subsequent pregnancy was 14% and 7% after singleton and twin pregnancies respectively

(Trogstad et al., 2004). A wide time interval between pregnancies and a change of partner are also associated with an increased recurrence risk of preeclampsia. The odds ratios (OR) for recurrence of preeclampsia in subsequent pregnancy are 11.8 (95% confidence intervals (CI) 11.1 - 12.6), and OR 8.2 (95% CI 5.9 - 11.3) with a different partner and the same partner respectively (Lie et al., 1998).

The clinical features of preeclampsia frequently resolve within days to a maximum of 3 months postpartum. However, it has been shown that preeclampsia is associated with long lasting morbidity for the mother, including an increased risk of type 2 diabetes (T2D), hypertension (HT), and coronary artery disease (CAD).

Libby et al (2007) examined the relationship between preeclampsia and the future development of T2D in their intergenerational cohort study composed of 7,187 mothers who give birth in Scotland between 1952 and 1958. They found that women with a history of preeclampsia were at increased risk of T2D in later life (OR 1.4, 95% CI 1.12-1.75, P= 0.003). This OR was adjusted for baby's birth weight and for social class (Libby, 2007). This was evident in a more recent study by Lykke et al (2009) who showed an increased risk for T2D in relation to hypertension in pregnancy. In this study, severe, recurrent, and early onset preeclampsia were associated with greater risk, after an average of 14 years follow up from the index pregnancy (Lykke et al., 2009). In agreement with this, in a study by Carr et al (2009) with a median follow-up of 8.2 years, the OR for diabetes in women with a history of preeclampsia was 1.82 (95% CI 1.26 -2.62) compared to women without preeclampsia after adjusting for age, parity, and gestational diabetes (Carr et al., 2009).

Persistence of metabolic syndrome and insulin resistance (IR) has been reported both in the early post-natal period after pregnancies complicated by preeclampsia and up to decades later in life. Sowers et al (1993) found altered plasma glucose and insulin responses to an oral glucose load in women 40 weeks after pregnancies complicated by preeclampsia, suggesting that an abnormality of carbohydrate metabolism persisted in those subjects (Sowers et al., 1993). Higher waist circumference, waist/hip ratio, body mass index, as well as increased serum insulin levels and lower glucose/insulin ratios have been reported in women with a history of preeclampsia compared to women with normal pregnancies (Pouta, 2004). Higher values of IR were reported in preeclamptic women one year after pregnancy compared to controls (Wolf et al., 2004). Higher levels of fasting insulin, lipids, coagulation factors, and adhesion molecules after pregnancies complicated by preeclampsia compared to controls have also been reported (He et al., 1999, Sattar et al., 2003). Girouard et al (2007) found significantly reduced levels of high-density lipoprotein (HDL) cholesterol and adiponectin, and elevated apolipoprotein B/apolipoprotein A1 ratio, homocysteine, leptin, and insulin levels among preeclampsia subjects compared with control subjects seven years after delivery (Girouard et al., 2007).

In relation to cardiovascular diseases, Ray et al (2005) analysed follow-up data from 1 million Caucasian women. They defined maternal placental syndrome as a composite of gestational hypertension, preeclampsia, placental abruption, and placental infarction, affecting 7.5% of the study population. They reported doubling of the risk of premature (before 38 years) cardiovascular disease in women who had had maternal placental syndrome compared to controls. The

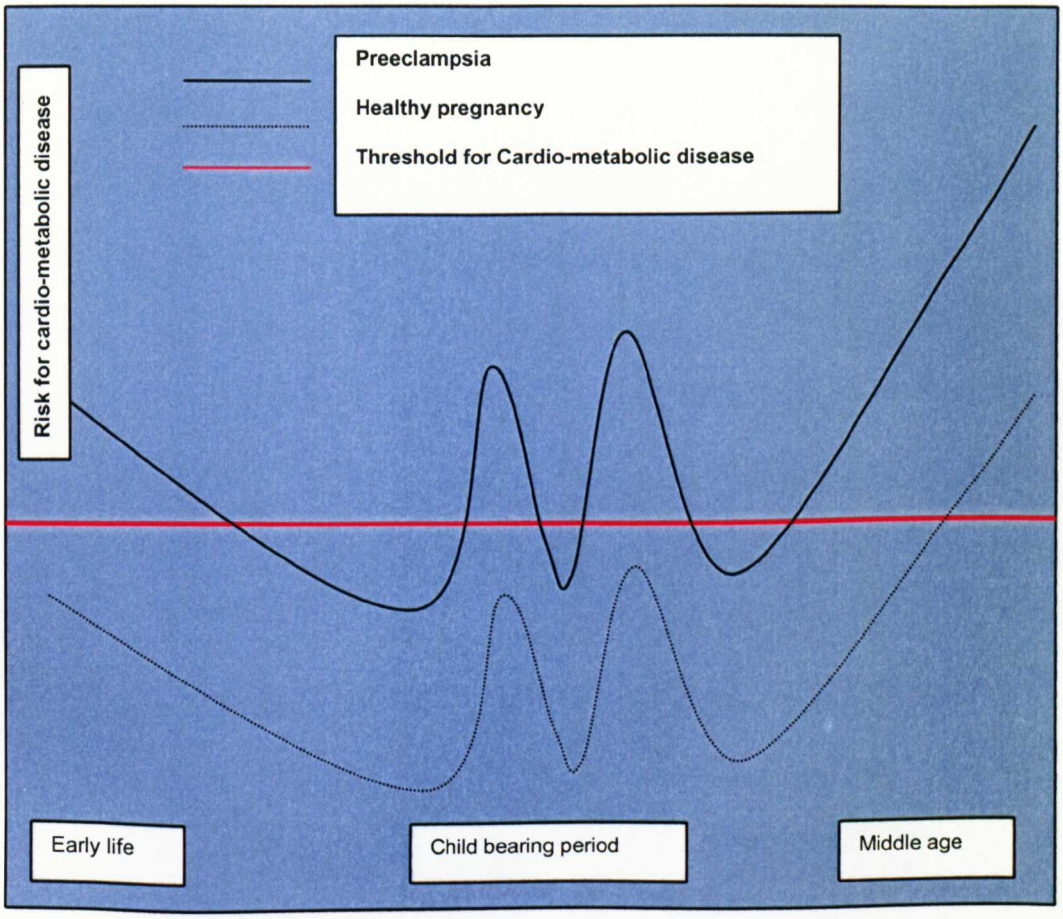
risk was tripled in the presence of fetal complications, or pre-existing risk factors for cardiovascular disease, including features of the metabolic syndrome and tobacco use. The association was evident when every component of the maternal placental syndrome was tested separately. The risk conferred by preeclampsia, which complicated 3.7% of cases, was comparable to that associated with the traditional risk factors for cardiovascular disease such as metabolic syndrome and tobacco use (Ray et al., 2005). This is in agreement with a more recent study carried out by Lykke et al (2009) of nearly the same size who reported that the risk for ischaemic heart disease was increased 1.48 fold (95% CI 1.25-1.76) after gestational hypertension, 1.57 fold (95% CI 1.44-1.72) after mild preeclampsia and 1.61 fold (95% CI 1.34-1.94) after severe preeclampsia following adjustment for maternal age, fetal birth weight, preterm delivery and T2D after 14 years follow up (Lykke et al., 2009). The increased effect associated with severe preeclampsia was confirmed in a study by Wikstrom et al (2005) in which the risk for ischaemic heart disease following mild preeclampsia was 1.9 (95% CI 1.6–2.2), compared with 2.8 (95% CI 2.2–3.7) following severe preeclampsia (Wikstrom et al., 2005). In a meta-analysis carried out by Bellamy et al (2007), the relative risk for hypertension was 3.70 (95% CI 2.70 to 5.05) after 14 years' follow-up, and 2.16 (95% CI 1.86 to 2.52) for ischaemic heart disease after 11.7 years (Bellamy et al., 2007). Levels of C reactive protein (CRP), an inflammatory marker of ischaemic heart disease, were found to be higher together with high blood pressure, an atherogenic lipid profile, higher fasting insulin and homeostasis model of insulin resistance in women 30 years after a diagnosis of eclampsia (Hubel et al., 2008). Although the studies described here were retrospective in nature with the possibility of recall bias for the study parameter, and the potential of unidentified confounders, the agreement among

studies on the effect of preeclampsia on future risk of maternal metabolic and vascular diseases, and the biological plausibility of the findings, are persuasive.

Pregnancy acts as a life long stress test for the mother. During pregnancy all maternal systems and organs must adjust to the increased demand of the growing fetus. Gestational diseases like preeclampsia develop when the mother is unable to cope with the increased metabolic and vascular load. Delivery is associated with remission, but there is evidence of residual cardiovascular compromise. This may indicate that preeclampsia induces irreversible metabolic and vascular changes in the mother rendering her susceptible to metabolic syndrome, T2D and CAD (Ramsay et al., 2003, Sattar and Greer, 2002), (Figure 1-1). An alternative explanation is that women with a predisposition to cardio-metabolic disorders are at increased risk of preeclampsia. Genetic risk factors may represent the link between preeclampsia and the cardio-metabolic disorders. A mother with cardiovascular risk factors including diabetes and chronic hypertension is currently considered as a high risk during pregnancy, and closely monitored accordingly. Similarly, it has been argued that women with history of preeclampsia would benefit from earlier screening for cardiovascular disease and life style modification in order to lower their risk of metabolic and ischaemic vascular disease.

Figure 1-1: Preeclampsia as a life long risk for cardio-metabolic disorders

Preeclampsia is the metabolic syndrome of pregnancy. Preeclampsia is associated with an increased risk of future cardiovascular and metabolic disorders (adapted from (Sattar and Greer, 2002).



1.1.2 Clinical manifestations of preeclampsia

The first criterion is a systolic blood pressure ≥140 mm Hg or a diastolic level ≥90 mm Hg on two or more occasions at least 4–6 h apart, at or after the 20th week of gestation. Clinically an increase in blood pressure of 30 mm Hg (systolic) or 15 mm Hg (diastolic) should prompt patient monitoring for preeclampsia. An increase in blood pressure is expected towards the end of an otherwise healthy pregnancy, so cannot be regarded as a pathognomonic sign of preeclampsia. The second diagnostic criterion is proteinuria, defined as ≥0.3 g protein in a 24 hour urine sample (Leeman and Fontaine, 2008, National High

Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy, 2000). When a 24 hour urine collection is not feasible, two random urine samples taken at least 4–6 hours apart can be used to measure protein/creatinine ratios; 30 mg protein/mmol urinary creatinine is regarded as the threshold for preeclampsia (Cote et al., 2008). The cheap, convenient urinary dipstick testing is widely used and ≥ +1 equivalent to ≥ 0.3 g protein provides a provisional diagnosis of proteinuria (Leeman and Fontaine, 2008). Results of dipstick testing should be interpreted with caution, because urine dipsticks can be affected by other maternal conditions e.g. dehydration (Sibai, 2003). Furthermore a negative test should not be ignored in a woman with new hypertension or symptoms or signs suggestive of preeclampsia; 12% of negative results will be false negatives as assessed against 24-hour urine collections (Magee et al., 2008).

The appearance of de novo proteinuric hypertension in pregnancy is diagnostic of preeclampsia, but some atypical presentations of preeclampsia occur. Hypertension after the 20th week of gestation without proteinuria can precede the development of the full spectrum of preeclampsia in 50% of cases. Persistent non-proteinuric hypertension is known as gestational hypertension; the severe form is typically associated with adverse maternal outcomes compared with mild preeclampsia (Buchbinder et al., 2002).

Pre-existing high blood pressure before pregnancy or before the 20th week of gestation is diagnosed as chronic hypertension, and up to 30% of women with this condition develop preeclampsia (Chappell et al., 2008). A sudden increase in blood pressure, new proteinuria, or signs and symptoms of other organ

involvement indicate superimposed preeclampsia. In the absence of pre-pregnancy records chronic hypertension may be overlooked due to the physiological fall in blood pressure during pregnancy. It is sometimes difficult to differentiate chronic hypertension from preeclampsia during pregnancy. Resolution of hypertension by 12 weeks postpartum suggests a retrospective diagnosis of preeclampsia. Persistence of hypertension beyond 12 weeks postpartum is often regarded as preeclampsia superimposed on chronic hypertension, although the possibility that massive endothelial damage caused by preeclampsia has resulted in chronic hypertension cannot be excluded. Similarly pre-existing proteinuria due to renal disease should be detected at early antenatal checks as this is associated with an increased risk of superimposed preeclampsia.

The preeclampsia syndrome may affect many body systems due to systemic endothelial activation and systemic inflammation, (e.g. impaired liver function, impaired renal function, coagulopathy, pulmonary oedema, and brain oedema) (Lowe et al., 2009). Consequently, women with new onset hypertension should be investigated with a full blood count, urea, creatinine, electrolytes, and liver function tests, and surveillance of fetal growth and amniotic fluid volume by ultrasound, and uterine and umbilical artery volocimetry should be carried out (Lowe et al., 2009).

Clinically preeclampsia is classified into mild and severe forms, depending on the clinical features. Severe preeclampsia is characterized by one or more of the following signs: systolic blood pressure more than 160 mm Hg and diastolic blood pressure greater than 110 mm Hg; proteinuria >5g /24 hours; oliguria < 500 ml in 24 hours; liver involvement manifested by elevated liver enzymes;

pulmonary oedema; thrombocytopenia; and involvement of the central nervous system in the form of convulsions (Leeman and Fontaine, 2008, National High Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy, 2000). HELLP syndrome, a triad of Haemolysis, Elevated Liver enzymes and Low Platelet count is a severe form of preeclampsia that can proceed to disseminated coagulopathy. Eclampsia, characterized by convulsions, complicates 2% of preeclampsia cases. 12% of cases of HELLP syndrome and 20% of eclampsia cases occur in normotensive nonproteinuric subjects (Landon et al., 2004, Magee et al., 2008, Steegers et al., 2010).

Preeclampsia is further classified into early onset and late onset depending on a time of diagnosis before or after 34 weeks of gestation (von Dadelszen et al., 2003). Early onset disease is less common (20%), more severe and usually complicated by the fetal syndrome of FGR and often associated with abnormal uterine and umbilical artery Doppler. Late onset preeclampsia is more common (80%), usually mild, and without fetal complication (Sibai et al., 2005). The differential diagnosis of hypertension in pregnancy is outlined in (Table 1-1).

Table 1-1: Classification and Diagnostic Features of Hypertension in Pregnancy

(Wagner et al., 2007), with permission.

Hypertensive disorders of pregnancy	Definitions and clinical manifestations
Preeclampsia/eclampsia	Preeclampsia, a pregnancy-specific disorder that occurs in 3%–5% of pregnancies, is a multisystem disease characterized by hypertension and protein level ≥300 mg in a 24-hour urine collection. Eclampsia, the convulsive form of preeclampsia, affects 0.1% of all pregnancies.
Gestational hypertension	Hypertension that occurs for the first time during the second half of pregnancy in the absence of proteinuria. It includes both women with preeclampsia who have not yet developed proteinuria and those with hypertension only; in a subset of patients with gestational hypertension, blood pressure remains elevated after delivery, leading to the diagnosis of chronic hypertension. Transient hypertension is characterized by blood pressure normalization by 12 weeks' postpartum.
Chronic hypertension	Blood pressure ≥ 140/90 mm Hg before pregnancy or before the 20th week of gestation. Most patients in this category will have a benign course with normalization of blood pressure midpregnancy.
Preeclampsia superimposed on chronic hypertension	Up to 30% of women with chronic hypertension develop preeclampsia, indicated by proteinuria that occurs for the first time in the third trimester, which is absent in uncomplicated chronic hypertension.

1.1.3 Risk factors for preeclampsia

A previous history of preeclampsia as well as a history of preeclampsia in a first degree relative (mother and sister) is reported to be associated with a doubling of the risk of preeclampsia (Nilsson et al., 2004, Plunkett et al., 2008). The partner's history is also important; being an offspring of a preeclamptic

pregnancy or fathering a previous preeclamptic pregnancy is considered a risk factor (Lie et al., 1998, Skjaerven et al., 2005). Based on these facts, a genetic component of preeclampsia including a role for paternal genes has been proposed. Preeclampsia is more common in nulliparous women (7.5%), and in parous women if they have wide inter-pregnancy intervals (>10 years), (Skjaerven et al., 2002) or have a new partner (Lie et al., 1998). Limited sperm exposure in these women might explain this increased risk, as a previous abortion or healthy pregnancy is associated with reduced preeclampsia risk (Trogstad et al., 2008).

Multiple pregnancy and hydatidiform mole, both associated with an increase in placental size, are also associated with increased risk for preeclampsia, indicating the central role of the placenta in its pathogenesis. Extremes of maternal age (< 20 and >40 years) also carry an increased risk of preeclampsia (Milne et al., 2005).

Pre-existing diabetes, high body mass index (BMI) before pregnancy, hypertension, renal disease, and thrombophilia are independent risk factors for preeclampsia. This is discussed in detail here because it is of particular interest in this study.

In 2003, Dunne et al carried out a study in the West Midlands, UK to investigate the effect of diabetes on pregnancy outcome and reported that the incidence of preeclampsia (among other adverse fetal and maternal outcomes) in 182 diabetic women was 20% compared with 10% in non-diabetic women in the local population (Dunne et al., 2003). Moreover, a study carried out by Rosenberg et

al (2005) in the US, including approximately 300,000 singleton pregnancies, reported that mothers with pre-gestational diabetes showed an increased risk for preeclampsia (OR 2.77; 95% CI 2.22-3.47). Gestational diabetes irrespective of ethnicity (White, Black, Asian, and Hispanic) was also associated with an increased risk of preeclampsia (Rosenberg et al., 2005). Moreover, a history of diabetes in a sibling was associated with an increased risk of preeclampsia in a study which included 190 cases and 373 controls (OR 4.7; 95% CI1.1 to 19.1) (Qiu et al., 2003).

Obesity is an independent risk factor for preeclampsia. It has been extensively studied in relation to preeclampsia, in which the risk escalates with different degrees of obesity. Compared with women with a BMI of 21, the adjusted risk of preeclampsia doubled at a BMI of 26 (OR 2.1, 95% CI 1.4 - 3.4) and nearly tripled at a BMI of 30 (OR 2.9, 95% CI 1.6 - 5.3). This was adjusted for diabetes, social status, smoking history and race and included 1179 nulliparous mothers (Bodnar et al., 2005). They showed also that women with a lower BMI (17 and 19) had a 57% and 33% reduction in preeclampsia risk respectively compared with women with a BMI of 21. In 2005, a larger study including 300,000 singleton pregnancies replicated this effect of BMI on preeclampsia risk (Rosenberg et al., 2005). It has been suggested that obesity confers risk for preeclampsia via increased insulin resistance as evidenced by higher leptin and lower adiponectin and sex hormone binding globulin (SHBG) in obese preeclamptic women compared to normotensive women (Catalano, 2010).

Hyperinsulinaemia early in pregnancy was able to predict the development of preeclampsia latter in pregnancy. IR as indicated by higher fasting insulin and

higher blood glucose and insulin levels after oral glucose measured in early pregnancy have been reported in several studies as a predictor for the development of hypertension in pregnancy (Solomon et al., 1999, Solomon et al., 1994). Low measures of SHBG, which correlates with insulin resistance, higher 24h urinary insulin as well as high HOMA-IR (Homeostasis Model of Assessment-Insulin Resistance) antedated the clinical diagnosis of preeclampsia (Emery et al., 2005, Sierra-Laguado et al., 2007, Thadhani et al., 2004).

This increased risk of preeclampsia in obese women, prediabetic, and diabetic mothers implies that these conditions might share the same pathological changes, including endothelial dysfunction, oxidative stress, systemic inflammation and coagulopathy. Conversely, as discussed earlier in this chapter, a history of preeclampsia is associated with an increased risk of cardio-metabolic diseases later in life. This indicates that certain shared risk factors may be present in early adult life, during the child-bearing period, and extending into middle age, making the affected mothers more vulnerable to both conditions (Figure 1-1). Genetic risk factors are strong candidates for the shared predisposition to preeclampsia and cardio-metabolic disorders in later life.

1.1.4 The role of the placenta

The aim of the placentation process is to anchor the fetus to the uterine wall, establish a means of communication between maternal and fetal circulations, and provide a barrier through which gas and nutrient exchange take place. This is accomplished by establishment of the placental villous tree and spiral artery

remodelling. Abnormalities involving these two processes are major characteristics of the preeclamptic placenta.

The central role of the placenta in the pathophysiology of preeclampsia is clearly evident, in that the presence of the placenta is essential for preeclampsia to occur and its delivery is the only cure. Preeclampsia is a complication of hydatidiform molar pregnancies which lack a fetus, and in women with postpartum preeclampsia curettage of the uterus leads to subsidence of preeclampsia symptoms (Young et al., 2010).

1.1.4.1 Insufficient spiral artery remodelling

Spiral artery remodelling is the transformation of the terminal branches of the uterine artery, occupying the endometrium and the inner part of the myometrium, from narrow coiled contractile high pressure vessels 200 μm in diameter into thin-walled flaccid low pressure vessels 2 mm in diameter (Figure 1-2). This ends with proper establishment of the feto-maternal circulation with an increase in the umbilical artery end diastolic velocity and an increase in the O_2 tension in the placental bed from 20 mm Hg to about 60 mm Hg. This process starts between the 8^{th} and 10^{th} weeks of gestation and becomes established around the 20^{th} - 24^{th} weeks of gestation during which there is an exponential increase in the fetal nutrient and gas requirement (Whitley and Cartwright, 2010).

During the early weeks of gestation, the highly proliferative villous cytotrophoblast (CT) cells stream out of the tips of villi and penetrate the overlying syncytiotrophoblast (ST) to form CT columns that invade the decidua

as extra-villous trophoblast (EVT). Two subtypes of the EVT are recognized, the interstitial and the endovascular. Interstitial EVT accumulates around the remodelling site where the decidual immune and stromal cells initiate the remodelling process (Lunghi et al., 2007). Endovascular EVT migrates into, and plugs, the terminal part of the vessel early in pregnancy to ensure the hypoxic state required at this stage. Later it erodes and replaces the endothelial cell (EC) lining of the vessel and the vascular smooth muscle cells (VSMC). Oxygen pressure must be maintained at defined levels throughout gestation so that the development of the early embryo during the first trimester of pregnancy takes place in a low oxygen environment (Jauniaux et al., 2000). Low levels of oxygen are important to ensure higher proliferative activity, lower invasive activity of CT, and reduced levels of reactive oxygen species (ROS), thus less DNA damage to the developing fetus (Huppertz et al., 2009).

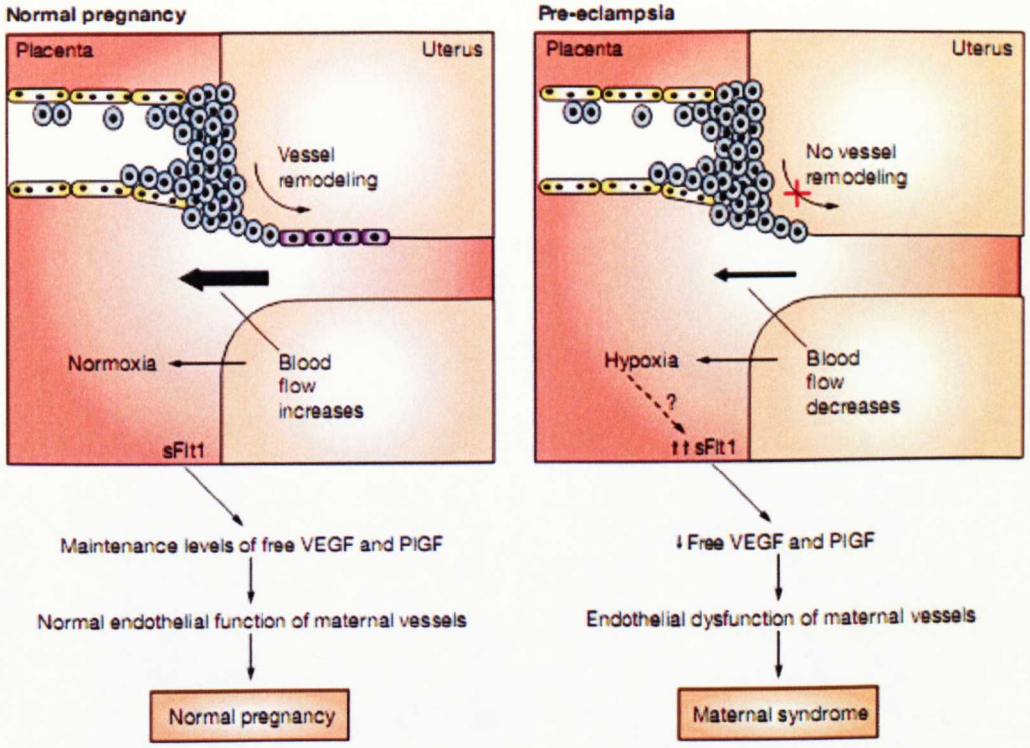
The process of remodelling involves apoptosis of the VSMC and EC as well as alteration in the extracellular matrix. A wide range of molecules is involved in this process. Tumour necrosis factor-α (TNFα), TNF-related apoptosis inducing ligand (TRAIL) and Fas ligand (FasL) are expressed by trophoblasts, and their receptors are expressed on both VSMC and EC with a shift in the balance of these molecules towards apoptosis (Harris et al., 2006, Ashton et al., 2005). Disruption of the extracellular matrix is thought to take place even before being in contact with the EVT as part of the decidualization process. Both trophoblast and VSMC produce matrix metalloproteinase(s) MMPs 1, 2, 9, and 12 that degrade the extracellular matrix with further disruption of the vessel wall (Harris et al., 2010).

In preeclampsia, invasion of the uterine spiral arteries is limited to the proximal decidua (Naicker et al., 2003). Both the invasive depth and numerical density of interstitial trophoblast are significantly reduced (Kaufmann et al., 2003), (Figure 1-2). This failure of trophoblast invasion in preeclampsia results in a reduction in utero-placental perfusion, with the placenta becoming increasingly ischaemic as gestation progresses. Reduced placental perfusion is evident as multiple central or large peripheral placental infarcts affecting 5% of the placenta, atherosclerotic changes in the placental vessels, endothelial damage, vascular thrombosis and fibrin deposition (Roberts and Post, 2008) as well as abnormal umbilical and uterine artery Doppler measurements (North et al., 1994).

Abnormal spiral artery remodelling leads not only to a reduction in the absolute amount of blood entering the placenta but also the manner by which the blood enters the inter-villous space. The blood enters the inter-villous space in an interrupted stream of jets of high velocity, as the vessel wall still retains its muscular layer. This creates a state of hypoxia /reperfusion that is deleterious to villous membrane integrity and function, and leads to enhanced apoptosis and necrosis. Furthermore, it will impose mechanical effects leading to excessive shedding of syncytial knots into the maternal circulation. It also leads to the formation of inter-villous lakes with the potential for thrombosis (James et al., 2010). The deformed chorionic plate and the increase in inter-villous space pressure also interfere with transport across the syncytial membrane. These changes in spiral artery remodelling occur primarily in early onset preeclampsia, often with FGR. Interestingly they may occur with isolated cases of FGR, indicating that such abnormalities may not be specific to preeclampsia (Huppertz, 2008).

Figure 1-2: Spiral artery remodelling in normal and hypertensive pregnancy

Defective spiral artery remodelling leads to placental hypoxia with subsequent release of soluble factors like sFlt-1 Into the maternal circulations leading to maternal systemic dysfunction and vasoconstriction (Redman and Sargent, 2005), with permission.



1.1.4.2 Abnormalities of the villous trophoblast

Under normal conditions, the villous tree starts to form upon implantation. It passes through three stages: primary, secondary and tertiary villi. The tertiary villi are formed of ST, a secretory multinucleated cell layer facing towards the maternal side of the villi. The second layer, facing towards the fetal side, is the CT. This layer is highly proliferative and responsible for continuously replenishing the ST layer by membrane–membrane fusion. The villi have also a connective tissue core that contains the newly formed fetal blood vessels. In the terminal branches of the villi, capillaries dilate into sinusoids covered by a vasculo-syncytial membrane (separating maternal and fetal circulations).

The vasculo-syncytial membrane consists of the ST and the endothelium of the capillary, separated by a joint basement membrane. The terminal villi are the physiologically most important components of the villous tree of the human placenta. The integrity of the ST is important because it carries a variety of transporters, receptors and enzymes that regulates maternal–fetal exchange (Jones et al., 2007).

Under normal conditions, ST is continuously replenished by fusion of CT cells, and the old and depleted ST undergoes apoptosis and shedding into maternal circulation as syncytial knots. The syncytial knots are cleared by the macrophages in the lung vasculature as they are too large to pass into the peripheral maternal circulation.

A large number of growth factors, adhesion molecules, cytokines, and cell cycle regulators have been implicated in the regulation of CT trophoblast proliferation and differentiation into syncytiotrophoblast, acting in autocrine and paracrine fashions. Many excellent reviews are available to describe such factors in details (James et al., 2010, Lunghi et al., 2007). Epidermal growth factor (EGF) and its receptors are expressed at the fetomaternal interface very early in pregnancy. EGF was found to regulate CT proliferation, differentiation and invasion in a gestational age dependent manner (Bass et al., 1994). It has been also implicated in abrogating the effect of apoptotic stimuli on the trophoblast cell lineage (Moll et al., 2007).

Vascular endothelial growth (VEGF) is one of the growth factors that has been extensively studied in relation to the placentation process; soluble receptors for this factor have been implicated in the pathology of preeclampsia (discussed in section 1.1.5). Both villous trophoblast and decidual immune cells are known to express VEGF and placental growth factor (PGF). VEGF expression is stimulated by hypoxia inducible factor-1 (HIF-1), under the hypoxic conditions encountered during early placentation. VEGF mediates its actions via 2 tyrosine kinase receptor isoforms, VEGF-receptor 1, (VEGF-R1 or flt-1) and VEGF-R2. VEGF-R2 binding by VEGF induces endothelial cell proliferation (associated with branching angiogenesis), while binding to VEGF-R1 stimulates endothelial tube formation. PGF is stimulated by oxygen and binds exclusively to VEGF-R1, resulting in non branching angiogenesis that dominates in the second half of pregnancy (Arroyo and Winn, 2008).

Placentas from early-onset preeclampsia exhibited a reduction in placental weights, volume of the intervillous space, surface area and volumes of the terminal villi. The picture was more extensive in the placentas from preeclampsia complicated by FGR, or in cases of FGR in which the intermediate and stem villi are also reduced in volume. The pathology involves reduction in the surface area of the villous capillaries and the blood vessels (Egbor et al., 2006a, Egbor et al., 2006b). Isolated cases of late onset preeclampsia are free of these changes (Egbor et al., 2006b, Mayhew et al., 2004).

The preeclamptic placenta is characterized by enhanced apoptosis and formation of more syncytial knots compared to normal placenta due to uncontrolled extrusion of apoptotic materials. It has been suggested these apoptotic materials

trigger systemic manifestations of preeclampsia (Gauster et al., 2009, Chen et al., 2010, Crocker et al., 2003). Syncytial knots are more common in placentas from preeclamptic pregnancies complicated by FGR compared to the normal villous histology of control pregnancies (Ishihara et al., 2002, Guller et al., 2008). Additionally, increased apoptosis may affect the transport function of the ST and consequently may be associated with FGR.

Moreover it has been noted that it is not only the apoptosis pathway which is enhanced; signs of aponecrosis and necrosis were reported in preeclampsia placentas. These pathways create subcellular particles of 100–200 nm known as the syncytio-trophoblast microparticles (STMPs) (Redman and Sargent, 2008). The ST trophoblast also releases smaller particles of 20–40 nm, termed nanoparticles or exosomes. Both STMPs and exosomes are rich in lipid and cytoplasmic components of the cell as well as DNA and microRNA of fetal origin. Consequently they are able to induce immune and inflammatory responses in the mother (Vanwijk et al., 2002, Redman and Sargent, 2008). These aponecrotic structures are released under normal conditions but in preeclampsia they are released in large numbers. In contrast to inefficient spiral artery remodelling and increased formation of syncytial knots, pathologies which are shared between preeclampsia and FGR, STMPs are only detected in preeclampsia (Goswami et al., 2006).

STMPs are placenta-derived microparticles, but other types of microparticles (MP) derived from platelets, granulocytes, T- cells and endothelium are also reported. These MP are suggested to be proinflammatory, procoagulant, and induce endothelial dysfunction in a group of diseases like CAD, diabetes, and

atherosclerosis (VanWijk et al., 2003). This suggests that MP might be a possible link between the local changes occurring in the placenta and the aberrant systemic maternal responses in preeclampsia.

1.1.5 Maternal systemic changes during pregnancy and preeclampsia

Aberrant maternal systemic responses in preeclampsia are the extreme of the adaptive mechanisms taking place during normal pregnancy (Roberts and Hubel, 2009).

Healthy pregnancy is associated with a state of insulin resistance, induced by hormones released by the developing placenta - placental growth factor, placental lactogen, progesterone and corticotrophin releasing hormone - to ensure the glucose supply to the fetus (Butte, 2000). An atherogenic lipid profile is created by mid-pregnancy (Hadden and McLaughlin, 2009).

Pregnancy is also characterized by enhanced coagulation activity due to elevations in plasma fibrinogen, von Willebrand factor, and other coagulation factors as well as increased platelet aggregability (Bremme, 2003). Increased inflammatory responses are also reported during normal pregnancy as indicated by leukocytosis (Sacks et al., 1998) and raised CRP compared to non pregnant states (Picklesimer et al., 2008).

Being responsible for a vast array of homeostatic functions such as blood pressure, blood flow, angiogenesis, coagulation, fibrinolysis, and inflammation, the endothelium is a major player in maternal adaptation for pregnancy.

Endothelial dysfunction is characteristic of preeclampsia, with an altered balance between vasodilator and vasoconstrictor mediators and an altered coagulation cascade secondary to endothelial damage. The damage is widespread and involves the placental vessels, renal, hepatic, cerebral and other systemic vessels. Soluble factors released secondary to placental hypoxia–reperfusion and STMPs are regarded as the link between abnormal placentation and widespread dysfunction of the maternal vascular endothelium. On the other hand, evidence that impaired endothelial functions precede preeclampsia suggests that endothelial dysfunction might be a cause, not an outcome in the course of preeclampsia (Granger et al., 2001).

The hypoxic placenta releases certain factors, many of which are known to interfere with angiogenesis and the integrity of the endothelium. Soluble fms-like tyrosine kinase-1 (sFlt-1) is an alternative splice variant of Flt-1 that lacks the transmembrane and cytoplasmic domains of Flt-1, and is released in large amounts into the blood of preeclamptic mothers compared to normotensive controls (Shibata et al., 2005, Levine et al., 2004, Gu et al., 2008). The high levels of sFlt-1 sequester both local and circulating VEGF and PGF and consequently interfere with their local and systemic effects. Locally it results in impaired angiogenesis, trophoblast proliferation and invasion. This was evidenced by an *in vitro* experiment showing defective human CT invasive capacity when sFlt1–Fc fusion protein was added to the culture medium. VEGF is important for integrity of the endothelium as well as the production of nitric oxide NO and prostacyclins, thus VEGF deficiency due to excess sFlt-1 results in systemic endothelial damage. Supporting evidence comes from experiments in

rats which manifested systemic vasoconstriction and proteinuria after administration of sFlt-1 (Maynard et al., 2003).

Endoglin is a co-receptor for transforming growth factor ß (TGF-ß) 1 and 3. The extracellular domain of endoglin can be shed as soluble endoglin (sEng), which is thought to impair TGF-ß1 binding to cell surface receptors and decrease endothelial nitric oxide signalling, thus inhibiting angiogenesis and promoting vascular dysfunction (Venkatesha et al., 2006) sFlt-1 and sEng are significantly increased in maternal serum prior to the onset of preeclampsia (Erez et al., 2008, Romero et al., 2008).

NO is also involved in impaired endothelial function in preeclampsia. NO enhances trophoblast functions (Dash et al., 2005). It is also responsible for systemic vascular dilatation occurring during pregnancy, and is thought to be a downstream mediator of VEGF; relative deficiency of this factor is associated with altered function of NO with subsequent loss of its vasodilator and antioxidant effects. Sandrim et al (2008) found that the level of nitrite, a metabolite of NO, is reduced in serum of preeclamptic women, correlating with higher levels of sFlt-1 and s-Eng (Sandrim et al., 2008).

The whole balance of ROS and antioxidant levels is challenged in preeclampsia. Many studies have reported an increase in maternal and placental production of ROS with subsequent lipid and protein oxidation (Serdar et al., 2003). Furthermore there is reduction in maternal antioxidant defence mechanisms as evidenced by reduced activity of superoxide dismutase and glutathione peroxidase, and reduced tissue and serum concentrations of antioxidant vitamins

(Aydin et al., 2004, Atamer et al., 2005). In spite of these suggestive observations, supplementation with vitamin C and E during pregnancy showed no effect on risk of preeclampsia (Rumbold et al., 2006, Poston et al., 2006).

An imbalance of anticoagulant and procoagulant processes is found in preeclampsia, as increased levels of proteins of the coagulation cascade have been reported in these women. Circulating levels of fibronectin are significantly increased as early as 20 weeks of pregnancy in women who subsequently develop preeclampsia, as well as other procoagulant proteins such as tissue plasminogen activator, plasminogen activator inhibitor-1, and fibronectin (Shaarawy and Didy, 1996). Von Willebrand factor and homocycsteine are elevated in preeclampsia while the levels of anticoagulant proteins like antithrombin III, protein C and Protein S are reduced (Kupferminc, 2003).

Dyslipidaemia is also a manifestation of preeclampsia, especially the early onset form (Clausen et al., 2001). It is characterized by increased low density lipoprotein (LDL), triglycerides, and free fatty acids and low high density lipoprotein (HDL) (Sattar et al., 1997, Kaaja et al., 1995). This is reflected in the placenta by a picture of acute atherosis due to uptake of the oxidized LDL by macrophages to form foam cells.

Oxidative stress, inflammation, dyslipidaemia, hypercoagulabilty and endothelial dysfunction are in a state of interaction and potentiate each other resulting in cumulative vascular and endothelial damage.

1.1.6 The relation between abnormal placentation and systemic maternal manifestations

It has been suggested that preeclampsia is a two stage disease, in which placental ischaemia induces and is followed by endothelial damage (Roberts and Hubel, 2009). The issue of a two stage disorder is questionable firstly because many cases, especially those which present with late onset preeclampsia, show no pathological features in the placenta. Secondly, the placental pathology appears in other pregnancy-related disorders such as preterm birth and isolated cases of FGR unrelated to preeclampsia. Lastly, markers of endothelial dysfunction (fibronectin, von Willebrand factor, adhesion molecules, reduced nitric oxide, and endothelin-1) have been reported in preeclamptic women weeks before the clinical presentation, indicating that endothelial dysfunction might not be a sequel of placental ischaemia but rather an independent contributor to the development of preeclampsia (Granger et al., 2001).

An oversimplified classification is: 1-) abnormal placentation combined with endothelial dysfunction leads to the early onset form of preeclampsia with FGR; 2-) abnormal placentation with normal maternal endothelium leads to isolated cases of FGR; 3) normal placentation with endothelial dysfunction leads to late onset preeclampsia. There is a possibility that both abnormal placentation and endothelial dysfunction share the same risk factors (Craici et al., 2008), which may be genetic in nature.

1.1.7 Genes and preeclampsia

1.1.7.1 Heritability of preeclampsia

A large body of evidence indicates that genes are risk factors for preeclampsia. There is a high recurrence rate of preeclampsia across generations (Arngrimsson et al., 1990) as well as an increase in the incidence of preeclampsia in first degree relatives of affected women. Daughters of mothers with a history of preeclampsia showed a 1.7 fold increase in the risk of preeclampsia; this risk was uniformly spread over first and second pregnancies, and was similar in daughters whether they were offspring of the preeclampsia pregnancy or not (Mogren et al., 1999). This implies that the risk is not related directly to the intrauterine environment but to inherited genetic factors. This was replicated in two other large studies from Sweden and Norway in which the mother to daughter risk ratio was estimated to be 2.6 and 2.2 respectively (Nielsen et al., 2003, Skjaerven et al., 2005). There is also an increased risk of preeclampsia among sibs as reported in different populations. Nilsson et al (2004) reported an OR of 3.3 (95% CI 3-3.6) for sibs of affected Swedish women. Similarly, a US study reported a sib OR of 4.52 (95% CI 4.21–4.83) and 4.11 (95% CI 3.59–4.63) in white and black women respectively (Plunkett et al., 2008).

The heritability of preeclampsia was estimated to be around 50%. Salonen Ros et al used the Swedish twin register to determine the role of genetic and environmental factors in the liability for preeclampsia and reported that monozygotic twins showed 57% concordance for preeclampsia compared to 18% in dizygotic twins, and estimated the heritability to be 54% (0–71%); the wide confidence interval was related to the small sample size (Salonen Ros et al.,

2000). Lower estimates of preeclampsia heritability of 30% were reported in more recent studies (Cnattingius et al., 2004, Nilsson et al., 2004). Cnattingius et al provided a detailed dissection of the genetic component of preeclampsia using the Swedish Birth and MultiGeneration Registries (Cnattingius et al., 2004). In his study, sister pairs, brother pairs and sister-brother pairs were used to differentiate between genetic and environmental components of preeclampsia and also to estimate the effect of maternal and fetal genes. The authors reported that 35% and 20% of the variance in liability of preeclampsia was attributable to maternal and fetal genetic effects respectively.

50% of fetal genes are maternally derived and 50% are paternally derived. Evidence for the involvement of paternal genes comes from the observation that the risk of fathering a preeclamptic pregnancy is increased among males who fathered a preeclamptic pregnancy with a different partner (Skjaerven et al., 2005, Lie et al., 1998) and among males who were themselves born from a pregnancy complicated by preeclampsia (Esplin et al., 2001, Sibai et al., 2005). Fetal genes are predicted to raise maternal blood pressure in order to enhance the utero-placental blood flow whereas maternal genes will act the opposite way. Endothelial dysfunction in preeclamptic mothers could, therefore, be interpreted as a fetal attempt to compensate for an inadequate utero-placental nutrient supply (Sibai et al., 2005).

Segregation analysis of preeclampsia pedigrees showed no clear pattern of inheritance. Early reports referred to a maternal recessive mode of inheritance as the incidence of preeclampsia is more common in mothers and sisters of the affected cases than in the general population, but is less than 50 % which is

expected in a dominant mode of inheritance, unless the genes were affected by low penetrance. However, recessive fetal inheritance was also proposed as the incidence in sisters-in-law is higher than expected in the general population (Cooper and Liston, 1979). Mitochondrial inheritance was suggested based on the finding that a mitochondrial gene defect was detected in preeclampsia (Folgero et al., 1996) but did not receive much support, partially because of the finding of a paternal contribution to the genetic risk of preeclampsia.

The nature of preeclampsia, being sex-limited and occurring only during pregnancy with the absence of known genetic carriers and the inconsistency of reports on consanguinity does not give solid support for the recessive mode of inheritance. In addition susceptibility to preeclampsia can be affected by maternal genotype, fetal genotype, or by an interaction of the two genotypes. Moreover, the contribution of the placenta, a major site for genomic imprinting, makes it difficult to assign a specific mode of inheritance for preeclampsia (Arngrimsson, 2005).

Currently preeclampsia is regarded as a complex genetic disorder in which many genes each contribute a modest effect to the phenotype. Susceptibility to preeclampsia is thought to result from complex interactions between maternal and fetal/paternal genotypes and environmental factors.

1.1.7.2 Genetic studies in preeclampsia

There are two main approaches to the identification of susceptibility genes: linkage and association studies. Each of these has 2 main components, either testing a candidate gene or chromosomal locus based on biological

understanding of the disease or searching the whole genome for loci that might affect the disease risk. In the following paragraphs the basic characteristics of each type of study and recent advances in relation to the genetics of preeclampsia will be discussed.

1.1.7.3 Genetic linkage studies

Linkage studies are based on the co-segregation of chromosomal loci with the disease in affected families. These chromosomal loci are identified by genotyping genetic markers assuming that these markers are co-inherited with the disease causing loci (individuals with similar phenotype share the same genetic markers). Linkage studies are effective in detecting risk variants that confer moderate to high disease risk. The markers used in linkage studies are usually microsatellite markers (polymorphic markers) so that segregation with the disease can be easily traced in families (Dawn Teare and Barrett, 2005).

Early studies of preeclampsia using the linkage approach focused on segregation of candidate loci in the affected family members. Linkage to the angiotensinogen gene on chromosome 1q42–43 (Arngrimsson et al., 1993) and to the endothelial nitric oxide synthase (eNOS) gene on chromosome 7q36 (Arngrimsson et al., 1997, DeFronzo et al., 1992) have been reported in Icelandic and Scottish families.

Later linkage studies did not focus on candidate loci but used a whole genome approach by genotyping evenly distributed markers along the whole genome in affected families. A number of chromosomal loci have been implicated in the risk of preeclampsia including 2p, 2q, 4p, 9q and 10q (summarized in

(Nejatizadeh et al., 2008). The chromosome 2p13 locus was first proposed by Arngrimsson et al (1999) in Icelandic families; linkage was related mainly to 2 multicase families. When these families were excluded from the analysis suggestive linkage appeared for a different locus on chromosome 2q23 (Arngrimsson et al., 1999). A subsequent genome-wide scan in families from Australia and New Zealand confirmed the presence of a susceptibility region on chromosome 2q22 located close to the loci identified in the former Icelandic study (Moses et al., 2000). Interestingly this chromosomal region was linked to essential hypertension in a UK family study (Padmanabhan et al., 2006). A case control study from Norway reported that an activin receptor gene is located in this region and 4 SNPs within this gene were associated with preeclampsia (Roten et al., 2009). Although these findings were not replicated in another study of Australian families (Fitzpatrick et al., 2009), a role for the activin receptor gene in preeclampsia can not be ruled out based on the potential biological effects of the gene in regulating decidualization and placentation.

A susceptibility locus on chromosome 9p13 was revealed by a genome-wide scan in 15 multiplex Finnish families (Laivuori et al., 2003). With respect to the association of preeclampsia with diabetes and cardiovascular disease mentioned earlier, it is of particular interest that the 9p13-p21 locus was shown to be a candidate region for type 2 diabetes mellitus in studies from Finland (Lindgren et al., 2002) and China (Luo et al., 2001). Recent genome wide association studies also demonstrated SNPs on chromosome 9p21 to be associated with T2D and CAD as will be discussed in the next chapter, making this region interesting for investigation. Harrison et al. (1997) found evidence of linkage to chromosome 4q. This region is close to the EGF gene region (Harrison et al., 1997), an

important regulator of trophoblast function and a candidate gene for FGR and preeclampsia.

The overall impact of the linkage studies is limited in defining the disease risk loci. Different study groups had their own inclusion and exclusion criteria, and identified different genetic risk loci. Ethnic heterogeneity (the disease co-segregates with different loci in different population) and phenotypic differences may have contributed to the lack of reproducibility of results. Linkage studies also have the inherent limitation of identifying only large chromosomal loci of around 30-50 cM that require fine mapping to locate the exact susceptibility region.

1.1.7.4 Candidate gene studies

Candidate gene association studies have been very widely used in the search for genetic variants which predispose to complex disorders. They test the prior hypothesis that biologically plausible genes or groups of genes in metabolic pathways are associated with disease risk by comparing the frequency of genetic markers, usually SNPs, in affected and unaffected individuals. The tested markers are mapped to the candidate genes and may purposely target putative functional SNPs. Compared to linkage studies association studies have more precision in identifying susceptibility markers that are relatively common in the population and confer a moderate to low risk of disease. Although this would improve understanding of the pathophysiological process of the disease, this approach is limited by the biological candidacy of the gene to previously suspected mechanisms.

Case control studies are the most common method used for genetic association studies. Cases and controls are easy to enrol compared to family-based studies so that large samples can be collected. Statistical analysis in the case of a dichotomous disease is easy to perform using Chi square testing and binary logistic regression to calculate the OR. Well defined phenotypes of cases and controls and adequate sample size are crucial to avoid spurious association. Careful selection of markers/SNPs is required to detect true association.

1.1.7.4.1 Selection of cases and controls and phenotype definition

Currently, the preeclampsia syndrome is viewed as an early-onset placental form (before 34 weeks of gestation), and a late-onset maternal form, mainly with endothelial dysfunction (after 34 weeks of gestation) (Redman and Sargent, 2005). These forms might be divergent in genetic risk and inheritance (Duckitt and Harrington, 2005). Late-onset, maternal preeclampsia arises from the interaction between a normal placenta and a constitutionally predisposed mother as occurs with pre-existing hypertension, diabetes or other states of metabolic imbalance associated with endothelial dysfunction. In contrast, early-onset placental preeclampsia starts with abnormal placentation, has a high recurrence risk, runs in families with a clear genetic component and is often associated with intrauterine growth restriction (Redman and Sargent, 2005). However recent studies have noted that endothelial dysfunction might cause abnormal placentation in early onset preeclampsia. This indicates that, although early and late onset preeclampsia may have different presentation, they share fixed risk factors, probably genetic in nature. This suggests that late onset–mild and early

onset-severe preeclampsia are parts of a continuum, and can be treated as one group for the purpose of genetic studies.

Any definition of preeclampsia will inevitably lead to the exclusion of some affected subjects with atypical presentations, for example those who develop gestational hypertension before developing proteinuria, women with chronic hypertension, T2D, obese women or multiparous. These criteria may actually select against those with the strongest genetic risk factors.

1.1.7.4.2 Selection of genetic markers for association study

Currently the most widely used markers for association studies are SNPs. This is because they are widely distributed across the genome (around one SNP in every 100-300 base pair (bp) and they are technically easy to genotype.

Currently around 11 million SNPs of minor allele frequency (MAF) >1% are estimated to be present in the human genome. The selection of which SNPs to test depends on many factors. Restricting analysis to SNPs with a MAF of $\geq$ 5% would be reasonable, as most sample sizes are inadequately powered to detect risk alleles with lower frequencies. Non synonymous SNPs, or SNPs mapped to regulatory regions (5'and 3' untranslated regions (UTR), promoter regions, or those in the putative splicing sites) are potentially functional and expected to be deleterious or beneficial in term of disease risk (Newton-Cheh and Hirschhorn, 2005). Recently data generated by the International HapMap project and the availability of high throughput low cost genotyping have facilitated comprehensive coverage of a whole gene or group of genes. This is

accomplished by selecting a group of tagging SNPs which provide a complete coverage of variation within the gene (s) of interest based on the presence of linkage disequilibrium (LD), (Gabriel et al., 2002). A pairwise linkage correlation coefficient (r^2) between two SNPs of 1 indicates that the markers are perfectly correlated, and can act as proxies for one another. This occurs when both markers are located within the same haplotype block, and greatly reduces the number of SNPs which must be genotyped to capture the underlying genetic variation.

As a result of LD, a genotyped marker which is associated with disease may be either the causal allele (direct association), or be in LD with the causal allele (indirect association). Association attributable to chance, selection bias, or other confounding factors may also occur (Cordell and Clayton, 2005).

When testing more than one SNP it becomes possible to test for the haplotype effect. Sometimes SNPs have small or no detectable effect on the disease risk when tested individually. When tested in cis they may show an effect, or a larger effect size, indicating either that there is an interaction between the tested SNPs or sometimes that the haplotype is acting as a proxy for non genotyped causal variants (Chapman et al., 2003, Palmer and Cardon, 2005).

1.1.7.4.3 Limitations of candidate gene association studies

Though association studies are a popular approach in mapping genetic variants for complex disease, they face the problem of lack of reproducibility. Association studies of preeclampsia reflect the situation in other complex

diseases. Goddard et al (2007) reported that of 127 polymorphisms in 61 genes tested for association with preeclampsia, 17 polymorphisms were significantly associated according to the authors of the studies. The association has been replicated for only nine polymorphisms in two or more studies (Goddard et al., 2007). The reason for this can be summarized into 3 major categories. Firstly a false positive association is correctly not replicated; secondly a true association fails to be replicated in an underpowered study, and lastly a true association in one population is not present in another due to variation in genetic and/or environmental background (Newton-Cheh and Hirschhorn, 2005).

Population stratification (admixture) is one reason for lack of replication. Population stratification occurs when populations with different ethnic origins and different rates of disease are mixed in a single study. If the genotyped marker is more common in one of these populations, a spurious association will result. So it is important in the design of any association study to match the cases and control with regard to ethnic background. Most studies depend on self-reported ethnic background (Freedman et al., 2004), but other strategies can be applied to avoid the effects of population stratification. The use of family-based association studies is one widely used solution. Transmission-disequilibrium testing (TDT) is the most popular method. The TDT design requires trios comprising an affected offspring and both parents. TDT depends on the biased transmission of the disease causing allele from heterozygous parents to affected offspring. The untransmitted allele acts as a control, so that any population-level allele frequency differences become irrelevant (Schaid, 1998). The disadvantage of TDT is the loss of statistical power because it relies on heterozygous parents only, which occur at a maximum of 50% for any given biallelic marker. It is

more laborious in terms of recruiting and genotyping trios, rather than individual cases and controls. It is also more sensitive to genotyping errors and missing data. Genome-based methods to correct for population stratification by genotyping random unlinked markers are also available. If true population stratification is present a disease-marker association test can be done in each matched subgroup; the total test for disease association is then a statistical combination of results from each component subgroup (Freedman et al., 2004).

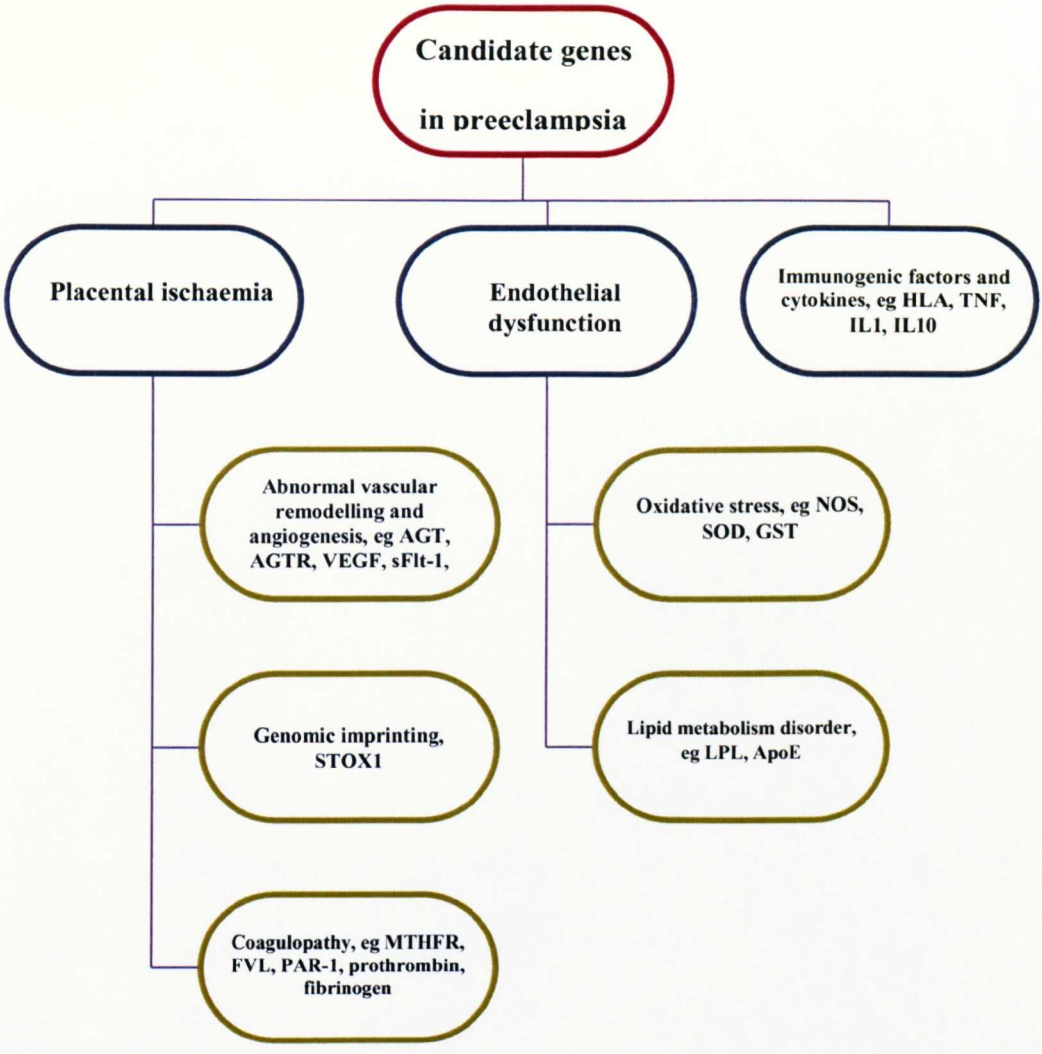
Inadequate sample size has a major impact on the ability to detect true association and on successful replication of association reported in previous studies. The threshold P value at which investigators report association is critical, especially when testing multiple markers and many phenotypes, when correction for multiple testing should be carried out (Newton-Cheh and Hirschhorn, 2005). A sample of several hundreds to several thousands (depending on the MAF) is needed to achieve a false positive rate of < 0.05. It is recommended that the power of the study should always be reported, and that the magnitude of the genetic effect which can be excluded is reported as the 95% confidence interval around the OR (Ioannidis et al., 2001). Genotyping errors, missing data, genotyping cases and control separately with variation in genotype call rate sometimes results in spurious association.

Additional factors which affect any genetic study include population and genetic heterogeneity, reduced penetrance, and phenocopies. Population heterogeneity is a condition in which the causative genetic variants are different among populations of different genetic and environmental backgrounds. Genetic heterogeneity refers to the condition in which different genetic variants lead to

the same phenotype. It includes locus heterogeneity, when various different genes lead to the same phenotypic expression and allelic heterogeneity, when more than one allele in a given gene is causative. Efforts to reduce genetic heterogeneity by using stringent definitions of phenotype or by testing subtypes of a disease can be considered. Reduced penetrance refers to the situation in which individuals carry the causative genetic variant without expressing the phenotype, and is the norm in genetically complex disorders. Conversely, a phenocopy is the situation in which an individual expresses the disease in the absence of any predisposing genotype. In complex disorders, reduced penetrance and phenocopies are frequently encountered, which may lead to misclassification of cases and controls, weakening the power of the study.

More than 60 genes have been studied in relation to preeclampsia. Candidate genes were chosen based on the pathways known to be involved in the pathophysiology of preeclampsia. They include genes related to vasoactive proteins/vascular remodelling, thrombophilia and hypofibrinolysis, oxidative stress, lipid metabolism, endothelial injury, placentation/imprinting and growth factors. Figure 1-3 highlights some of these genes (Mutze et al., 2008). There is lack of replication in most of the studies due to the aforementioned factors and. no single gene has been unequivocally established as a susceptibility gene for preeclampsia (Chappell and Morgan, 2006).

Figure 1-3: Some of the candidate gene studied in preeclampsia



Two impressively large association studies in preeclampsia are worth noting because both of them addressed many of the potential limitations mentioned earlier, and they reflect the current situation of the genetics of preeclampsia. In addition, both of these studies investigated the role of maternal and fetal genes on the risk of preeclampsia. The first was carried out in the UK as a British multicentre candidate gene study, by the GOPEC consortium. In this study 28 SNPs in 7 genes (angiotensinogen, (AGT), angiotensin receptor type 1 (AGTR1) and 2 (AGTR2), TNF-α, endothelial nitric oxide synthase (NOS3), factor-V Leiden, and methylene tetrahydrofolate reductase (MTHFR)), commonly

reported to confer susceptibility for preeclampsia, were investigated. TDT was applied to 398 preeclampsia maternal (mother and her parents), and 536 fetal (preeclampsia mother, her partner and the offspring) trios. Besides avoiding population stratification by using TDT test, the study also applied a stringent level of significance by testing the association in the context of prior probability of gene candidacy and reported an 85% power to detect modest genetic risk. Of the 28 SNPs that had been investigated only 2 markers in AGTR2 (fetal and maternal) one marker in AGT and TNF-α for fetal genes showed a marginal protective effect (genotype relative risk range from 0.48-1.0), excluding a major risk conferred by these genes (The GOPEC Consortium, 2005).

The second study was conducted by Goddard et al (2007) who investigated 775 SNPs in 190 candidate genes in 394 preeclamptic women and their 342 offspring in comparison to 602 healthy pregnant control women and their 631 offspring. Again the common candidate genes MTHFR and NOS3 were not significantly associated with preeclampsia. Maternal collagen type I alpha 1 (COL1A1) and interleukin I alpha (IL1A) genotype, fetal plasminogen activator, and urokinase receptor genotype, and fetal and maternal insulin like growth factor 1 (IGF1) and interleukin 4 receptor (IL4R) gave promising results. The study had some limitations; including the sample size in relation to the number of SNPs investigated which is reflected in the lack of significance for all variants after correcting for multiple testing. They also used whole genome amplified DNA rather than genomic DNA with the possibility of sequence errors. The authors reported the study had 80% power to detect a genotype relative risk of between 1.6 and 1.8, so smaller effects might remain undetected (Goddard et al., 2007).

1.1.7.4.4 Awaited genome wide association studies (GWAS)

The completion of the Human Genome Project and of the International HapMap project has increased knowledge of the structure of the genome. Massive advances in high throughput genotyping technology have enabled the genotyping of up to 1 million genetic variations, (mainly SNPs) in a single experiment using SNP chips. The advances have made GWAS possible. GWAS is an unbiased approach with the potential to detect new causative variants and shed light on novel biological mechanisms that may help in understanding and managing the disease under investigation. For such studies to achieve statistical power thousands of cases and controls are required and replication of significant results in an independent sample set is essential. The first GWAS for preeclampsia is underway now, as a part of the Wellcome Trust Case Control Consortium 3 (WTCCC3).

1.2 Fetal growth restriction

Fetal growth restriction (FGR) is a condition in which a fetus fails to achieve its full growth potential due to reduced tissue deposition and/or reduced nutritional supply from the utero-placental circulation (Cetin et al., 2004). FGR is a complex disorder affecting up to 10% of neonates.

Normal fetal growth passes through different stages. The early stage involves the proliferation, organization, and differentiation of the embryo (embryonic life). The later stage comprises continuous growth and functional maturation of the different tissues and organs of the fetus (fetal life). The acquisition of fetal weight progressively increases with advancement of gestation from 5 g/day

during the 1st trimester and reaches around 30–35 g/day at 34 weeks. Therefore, most fetal weight gain occurs in the last 20 weeks of gestation (Monk and Moore, 2004).

By mid-gestation two significant events have occurred: establishment of the fetomaternal circulation, and secretion of measurable quantities of insulin by the fetal pancreas. This ensures patent maternal supply lines to, and resource utilization by, the growing fetus. Glucose is the principle source of energy for the fetus. Its constant availability and supply by the mother, placental transport by specialized adaptive mechanisms, along with fetal insulin secretion facilitate glucose utilisation by the fetus (Hay, 2006).

Securing adequate nutrient and oxygen supply to the developing fetus depends on the genetic profile of the embryo (maternal and paternal genes), the maternal–placental–fetal unit, and environmental factors (hormones, nutrition, infection, smoking) in the maternal milieu (Murphy et al., 2006).

1.2.1 Definition of FGR

Classification of birth weight has been based on population growth curves after correcting for gestational age. Babies were classified according to their birth weight centile (BWC) into 3 categories: infants with birth weights<10th centile, called small for gestational age (SGA); infants with birth weights between 10th and 90th centile, called appropriate for gestational age (AGA); and infants with birth weights >90th centile, called large for gestational age (LGA) (Reeves and Bernstein, 2008b).

The problem with this classification is that a subset of SGA and LGA babies are simply constitutionally small or large due to, for example, a small or large sized mother, or representing the tail of the normal distribution. On the other hand some babies with a BWC >10 manifest some features of growth restriction, because the definition of SGA does not address true intrauterine growth but rather size at birth (Gardosi, 2004). So, SGA is an inadequate proxy for fetal growth restriction (FGR). The most commonly used definition of FGR is a sonographically estimated fetal weight (EFW) that falls below the 10^{th} centile for gestational age (American College of Obstetricians and Gynecologists. 2000).

Ideally, SGA definitions based on population derived birth weight curves would utilize data from healthy pregnancies only (Reeves and Bernstein, 2008b, Goldenberg and Cliver, 1997). Birth weight curves adjusted for gestational age alone ignore some of the biological factors that might affect fetal or birth weight variations, including fetal sex (girls are lighter than boys), ethnicity (European are heavier than East Asian populations), parity (which is positively correlated with birth weight), and maternal height and weight. These factors were incorporated into customized antenatal charts (Gabriel et al., 2002). With the use of the customized charts, 28% and 22% of SGA and LGA babies from conventional charts were recategorised as AGA.

Important criteria are required for reliable estimation of the fetal weight centile, including: 1. an accurate gestational age estimation, using ultrasonic measurement of the crown rump length in the first trimester pregnancy, the head circumference in second trimester, or the date of the last menstrual period; and 2. an estimated fetal weight using at least 3 variables: biparietal diameter,

abdominal circumference and femur length. A weight percentile can then be calculated from the estimated weight and gestational age using available curves (Platz and Newman, 2008).

The estimation of fetal weight centile has some limitations. Ultrasonic measurements are technically difficult and subject to inaccuracy due to inter and intra observer variability. Population data derived from a large number of pregnancies are essential for reliable estimates of fetal weight centile. Around 25% error between the actual and predicted weights can occur, particularly in the estimation of macrosomia or FGR, most probably due to limitations in estimating soft tissue mass. Incorporation of other indices like amniotic fluid volumes and umbilical artery Doppler velocimetry has been shown to improve the diagnosis of FGR (Platz and Newman, 2008, Gardosi, 2004).

Gardosi et al (1995) incorporated the factors shown to affect variation in birth weight (gestational age, fetal sex, ethnicity, maternal weight, height, and parity) into a computer software package from which the expected term birth weight can be calculated as the customised birth weight centile or term optimal weight (TOW). To further adjust for intrauterine growth, estimated fetal weight based on ultrasonic measurement between 24 and 42 weeks of gestation expressed as proportion of the term birth weight was incorporated in this software, termed gestation related optimal weight (GROW). From this software the birth weight can give insight into fetal growth (Gardosi et al., 1995). These customized curves are more accurate in predicting the neonatal outcome, low Apgar score, neonatal intensive care admission and neonatal death (Gardosi, 2004, Gardosi et al., 2009). When applied retrospectively to mothers who had induction of labour due

to FGR based on population curves, the customized curves suggested that 60% of these pregnancies were free of FGR (Avery et al., 2006). They were also more successful in identifying SGA babies who were truly growth restricted and more prone to later metabolic disturbances (Verkauskiene et al., 2008). The customized curves are valuable in identifying growth restriction in preterm babies, who form a large proportion of the SGA babies (Ananth and Vintzileos, 2009). It cannot be applied to twin pregnancies.

It is clear that not all FGR babies are SGA and vice versa. Nevertheless, SGA is still used as a surrogate for FGR especially in the research environment. Some variation exists in the cut-off values at which the diagnosis of FGR is made. This ranges between $<3^{rd}$ up to 15^{th} centile. The use of the 3^{rd} centile was associated with a higher OR for adverse neonatal outcomes (100-fold increase), but it also missed many FGR cases. On the other hand, although a large proportion of the $<10^{th}$ centile group turn to be constitutionally small, the use of this cut off threshold can still detect up to a 30-fold increase in adverse neonatal outcomes and was found to be highly correlated with still birth related to SGA (Platz and Newman, 2008).

1.2.2 Implication of FGR

FGR complicates up to 10% of all pregnancies. It is therefore a common obstetrical condition, but hard to manage due to its high rate of complications and adverse consequences. A history of SGA in a previous pregnancy was found to increase the risk of SGA in a subsequent pregnancy (OR 2.5; 95% CI 1.7-3.3); the risk was greater when combined with smoking and low maternal BMI (Bakketeig et al., 1993). FGR complicates about 30% of preeclamptic

pregnancies and is regarded by some as a sign of severe preeclampsia. It also represents an increased risk for preeclampsia in a subsequent pregnancy (Lie et al., 1998).

The complications of FGR are either immediate in the perinatal period or occur later in adult life. Immediate complications include hypothermia, hypoglycaemia, pulmonary haemorrhage, encephalopathy, necrotizing enterocolitis, and retinopathy, (McIntire et al., 1999, Benedict et al., 2001). Neonatal mortality is high in the SGA babies especially in those who are preterm, who have a 33% mortality compared to 17% in AGA preterm babies (Giapros et al., 2011).

A large body of epidemiological data has accumulated suggesting that being small sized at birth is associated with a substantial risk of developing adult diabetes, hypertension, and coronary heart disease especially when combined with catch-up growth during childhood (Barker et al., 1989a, Barker et al., 2002, Osmond and Barker, 2000).

A study including around 15000 subjects born from 1911 to 1930 in Hertfordshire, UK, showed that death rates from CAD increase in relation to low birth weight and are also independently affected by lower weight estimates at the age of 1 year (Barker et al., 1989b). These data have been confirmed in different populations including the Finnish population, using the Helsinki birth cohort study of nearly the same sample size as the Hertfordshire study (Eriksson, 2007) and a study including 70000 women from the US (Rich-Edwards et al., 1997). A combination of low birth weight and high BMI at the age of 19 was associated

with an increased risk of CAD in Danish subjects after adjustment for social and educational class (Osler et al., 2009).

Barker et al (1993) showed that individuals with low birth weight have a 10-fold increased risk of developing metabolic syndrome at the age of 64 compared to those who have a higher birth weight, and that this was independent of prematurity or smoking habits (Barker et al., 1993). T2D was also associated with low birth weight in the Hertfordshire collection. The OR was 1.3 with low birth weight and low BMI and increased to 2.5 when the BMI was higher at the age of 11 years (Barker, 2002). This was replicated in Swedish (Lithell et al., 1996) and Taiwanese populations (Wei et al., 2003). Thinness at birth was also associated with insulin resistance (Jensen et al., 2002), and markers of endothelial dysfunction (Pellanda et al., 2009). Moreover parents of SGA infant are at increased risk of T2D and insulin resistance (Lindsay et al., 2000, Lawlor et al., 2002).

1.2.3 Aetiology and risk factors for FGR

It has long been appreciated that FGR is not a single disorder, but instead has various causes. A list of factors associated with increased risk for FGR is presented in Table 1-2 (reviewed in (Hendrix and Berghella, 2008, Murphy et al., 2006)).

Table 1-2: Fetal and maternal risk factors for FGR pregnancy

(adapted from (Hendrix and Berghella, 2008).

Fetal
• Genetic diseases (5% to 20%) (chromosomal abnormalities, single gene defect, and polygenic disorders)
• Infection (5% to 10%) (toxoplasmosis in developed and malaria in developing countries)
• Malformations (1% to 2%)
• Multiple gestation (3%)
• Placental/cord abnormalities (3%) (abruption, mosaicism, 2 vessel cord, placenta praevia, abnormal umbilical artery insertion, etc.)
Maternal
• Hypertensive disorders (20% to 30%)
• Pregestational diabetes
• Autoimmune disease (systemic lupus erythematosus)
• Cardiac disease (eg, complex cyanotic congenital heart diseases)
• Toxin exposure (smoking, alcohol, cocaine)
• Malnutrition
• Living at high altitudes
• Living in developing country
• Low socio-economic status
• Race (eg, African-American)
• Family or prior history of FGR
• Extremes of maternal age
• Infertility/assisted reproduction technology
• Short inter-pregnancy interval

Genetic factors, notably chromosomal aberrations like trisomy 21 (Downs syndrome), trisomy 13, and trisomy 18 are known to exhibit FGR mainly due to abnormal development of the placental vessels. Deletion of part of a chromosome can also lead to FGR as is the case with chromosome 4q known as Wolf-Hirschhorn syndrome. Confined placental mosaicism and uniparental

disomy (e.g. UDP7) can also lead to FGR (reviewed in ((Hendrix and Berghella, 2008).

Toxoplasmosis and cytomegalovirus in developed countries and malaria in developing countries are maternally transmitted infections that affect growth and development of the fetus by affecting the vascular system and by causing congenital malformations. It has recently been suggested that H. pylori infection in pregnant women may affect fetal intrauterine growth (Eslick et al., 2002).

Preeclampsia is a major risk factor for FGR, as around 30% of preeclampsia is associated with FGR. Both conditions have shared pathology in the form of impaired trophoblast invasion and apoptosis and limited villous development. Underlying maternal medical conditions which lead to reduced oxygen carrying capacity (anaemia, high altitude) or affect the maternal vascular system (sickle cell anaemia, systemic lupus erythematosis, anti-phospholipid syndrome, and thrombophilias) are also associated with FGR pregnancies (Chakravarty et al., 2008).

Maternal diabetes, hyperglycaemia, or insulin resistance have long been known to affect birth weight in the direction of macrosomia, resulting in LGA babies (Ong et al., 2008). This is due to increased fetoplacental availability of nutrients in late gestation due to maternal hyperglycaemia (Hay, 2006). Diabetes also is reported to cause SGA especially if it is long-standing diabetes associated with vascular injury (Haeri et al., 2008, Howarth et al., 2007). A U-shaped relationship between maternal diabetes and birth weight was described. In fact

both SGA and LGA in diabetic pregnancy are subject to adverse fetal and long lasting outcomes.

Low maternal BMI (BMI <20) is associated with an increased risk of having a low birth weight baby, whilst maternal BMI > 25 is associated with high birth weight (Baeten et al., 2001, Bhattacharya et al., 2007).

Smoking is a major (avoidable) risk factor for FGR. Mothers who smoke 4-5 cigarettes per day have a 2-3 fold increased risk for SGA, with a reduction of 150-300g in birth weight (Wang et al., 2002, Aagaard-Tillery et al., 2008, Dejmek et al., 2002). Smoking increases carbon monoxide, reduces NO and causes vasoconstriction, limiting the oxygen supply to the fetus (Reeves and Bernstein, 2008a). It is also suggested that smoking reduces mitochondrial respiratory function in the placenta and consequently reduces energy availability (Bouhours-Nouet et al., 2005). Smoking is associated with a reduced risk for preeclampsia (England et al., 2002, Mortensen et al., 2001, Broughton-Pipkin, 2008) although smoking causes endothelial dysfunction.

1.2.4 The role of the placenta in FGR

The main determinants of fetal growth can be categorized into 3 major components: fetal, placental, and maternal. These factors control the genetic profile of the embryo, proper placentation, the integrity of the materno-placento-fetal unit, adequate nutrient and oxygen supply to the developing fetus, and the hormonal milieu. These factors typically interact with each other to optimize the outcome for the baby and the mother.

Shallow trophoblast invasion and abnormal spiral artery remodelling, increased apoptosis of the villous trophoblast, and reduced vascular development, are shared pathological features between preeclampsia and FGR as described earlier (Section 1.1.4). A state of hypoxia develops secondary to impaired development of the placenta.

However specific changes in placental transport properties are described in the FGR placenta. Hypoglycaemia and reduced essential amino acid concentrations were reported in SGA babies. *In vitro* studies using isolated microvillous membrane suggested that FGR is associated with a number of changes in the activity and/or expression of placental nutrient and ion transporters. For example Cetin et al (2002) reported that system A (for essential amino acids, (AA)) and system L (for arginine amino acid) transporters were reduced in placenta from FGR fetuses, whilst they maintain glucose transporters. Transfer of polyunsaturated fatty acids from the maternal circulation across the placenta was found to be disturbed in FGR as reflected by lower fetal / maternal ratio for these fatty acids in FGR compared to control babies during the second half of pregnancy (Cetin et al., 2002). Modifying the placental delivery of fatty acids to the fetus in a crucial period when fat deposition increases exponentially and is required for neural development may explain neural defects in FGR babies (Cetin et al., 2002).

The placenta also has an important endocrine function. It is responsible for the release of many hormones including estrogens, progesterone, human placental lactogen (hPL) and human growth hormone splice variants (hGHV). These hormones create a state of insulin resistance in the mother and ensure adequate

glucose and amino acid supply to the growing fetus. The level of hPL in the mother was found to be correlated with fetal growth velocity (Sorensen et al., 2000). Lower maternal levels of hPL were reported with FGR (Bersinger and Odegard, 2004).

1.2.5 Genetics of FGR

The influence of genes on birth weight is reflected in lower birth weight in relation to female sex. The heritability of birth weight has been estimated as between 30% and 90%. In a study of mother-father-baby trios from Norway, maternal and paternal birth weight were independent predictors of offspring birth weight, with the coefficient of correlation between baby's birth weight and maternal and paternal birth weight calculated as 0.22 and 0.13 respectively. The heritability estimate was 25% in this study (Magnus et al., 2001). The effect of maternal and paternal weight on birth weight was replicated in a UK population (Griffiths et al., 2007). Higher effects of maternal birth weight on the offspring birth weight may be explained by both genetic effects and the intrauterine environment.

In a more recent study a high heritability estimate for birth weight (92%) was reported in a Hispanic population, adjusted for gender, gestational age, birth order, and a number of maternal factors including age or gestational diabetes (Cai et al., 2007). It should be noted that heritability estimates are population-specific, and can be influenced by such factors as the environment, study population, ascertainment criteria and the covariates considered in the analysis. A high degree of heritability can be achieved by adjusting for covariates that explain some of the environmental factors in the analysis.

It is not only the variation in normal birth weight which is thought to be heritable but also the condition of low birth weight or SGA. It has been shown that the infant birth weight pattern tends to be repeated in subsequent pregnancies (Fabio Ghezzi, 2003, Magnus et al., 1997). In a study which included 3,505 French sibs from term singleton pregnancies, La Batide-Alanore et al. (2002) reported sib-sib odds ratios for SGA to be 4.8 (95% CI: 3.7, 6.3). The findings were significant when SGA was defined as $<10^{th}$ centile as well as $<5^{th}$ centile corrected for gestational age and baby's sex (La Batide-Alanore et al., 2002). This was also reported in a recent study in the US which showed that the risk of being SGA is 25% when the older sibling was SGA compared to only 6% when the older sibling was AGA (Ananth et al., 2009). This familial aggregation of SGA indicated a genetic element as it remained after adjustment for maternal and environmental factors.

Moreover, women who were themselves FGR are more likely to deliver an FGR baby (Leff et al., 1992, Winkvist et al., 1998). A paternal history of being SGA was associated with doubling of the risk of fathering a SGA baby (Magnus et al., 1997).

1.2.6 Genetic studies of FGR

It has been assumed that the genetic effects on birth weight are complex in nature with many independently loci acting on the phenotype in an additive fashion that can be modified by environmental factors (Lunde et al., 2007). Several studies have been performed to examine the association between genetic variants in biologically plausible candidate genes and FGR.

1.2.6.1 Fetal programming and diabetes genes

Two theories have been described to explain the relation between low birth weight and adult disease. The first was the fetal programming or Barker hypothesis proposed two decades ago. Fetal programming postulates that a malnourished fetus due to an adverse intrauterine environment is programmed to exhibit a thrifty phenotype with increased food intake and fat deposition. This is to ensure survival of the fetus in the short term, for the duration of pregnancy. But when the situation changes after birth and the newborn is provided with abundant nutrients, such individuals develop obesity and T2D as adults (Barker, 1995, Barker, 2002, Barker et al., 2002). This is consistent with the observation that these disorders are mainly manifested in low birth weight individuals who experienced catch up growth in early childhood, not in those children who remained lean. Excess glucocorticoids described in low birth weight babies have also been implicated in later organ and system dysfunction. Excess glucocorticoids lead to reduced beta cell mass in the pancreas and reduced insulin secretion, increased hepatic glucose production, reduced hepatic glucose uptake, increase lipolysis and increased insulin resistance of the skeletal muscles, ultimately leading to glucose intolerance and T2D (Ross and Beall, 2008).

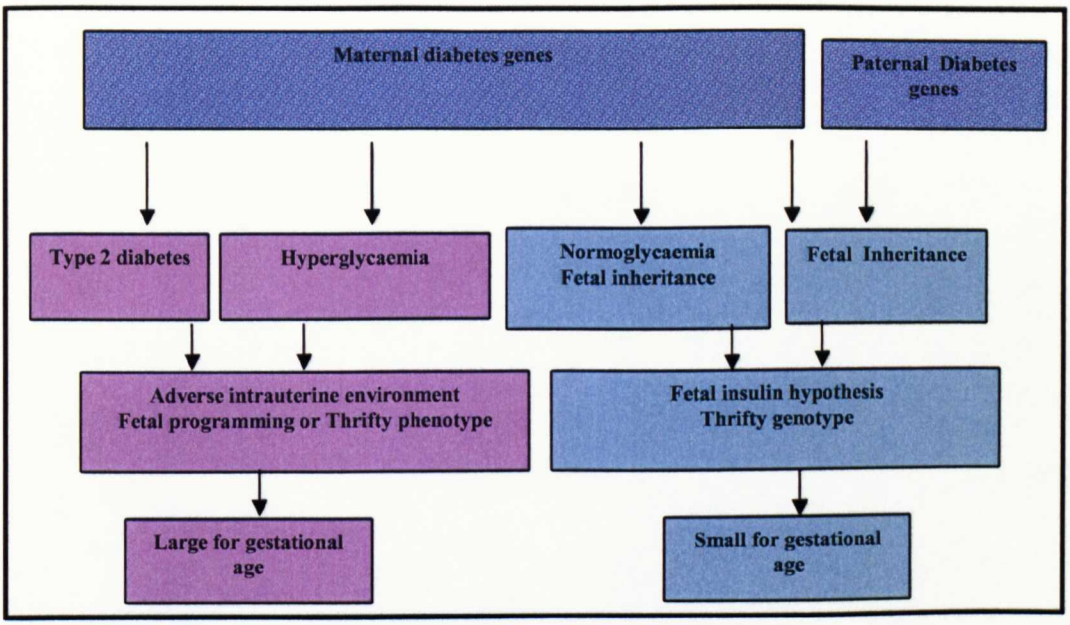
Animal studies indicated that epigenetic changes in the form of DNA methylation or histone modification which alter gene expression without changes in DNA sequence might explain fetal programming. For example in a rat model of FGR, methylation of pancreatic duodenal homeobox-1 promoter was associated with reduced insulin secretion (Park et al., 2008).

In parallel with fetal programming, the fetal insulin hypothesis has been advanced to explain the mutual association between T2D and low birth weight. The fetal insulin hypothesis, or thrifty genotype, suggests that the same genes which are responsible for low birth weight due to low insulin secretion and/or insulin sensitivity lead to T2D later in life (Frayling and Hattersley, 2001). The association between T2D in parents and the birth weight of their offspring gave strong evidence for the role of diabetogenic genes in controlling birth weight. Lindsay et al (2000) reported an association between paternal diabetes and low birth weight; low birth weight in offspring conferred a 1.8-fold increased risk on the father of developing diabetes later in life (95% CI 1.2-2.7) (Lindsay et al., 2000).

It is widely recognised that diabetic mothers commonly give birth to LGA babies. However, the role of maternal diabetes susceptibility genes on the offspring birth weight was suggested by two studies which showed that low birth babies were born to mothers who were not diabetic at the time of pregnancy but became diabetic or insulin resistant later in life (Hypponen et al., 2003, Lawlor et al., 2002). These findings lend support to the fetal insulin hypothesis or thrifty genotype, and are consistent with the fetus inheriting 50% of his/ her genes from the father and 50% from the mother which both predispose to diabetes and reduced fetal growth (Figure 1-4).

Figure 1-4: The effect of maternal and paternal diabetes genes on birth weight

Fetal and maternal genes are interacting in relation to offspring birth weight, the fetal programming and fetal insulin hypotheses are not mutually exclusive.



A rare mutation in the fetal glucokinase gene (GCK), a diabetes susceptibility gene, was associated with a 533 g reduction in birth weight and mild post natal hyperglycemia. The effect of maternal genotype was in the opposite direction with a 601 g increase in the offspring birth weight. When both mother and fetus carried the mutation the effect on birth weight was nil (Hattersley et al., 1998). These mutations could cause impaired fetal insulin secretion, hence their contributions to impaired fetal growth and to subsequent diabetes, whilst maternal effects were mainly caused by the intrauterine environment associated with hyperglycemia. Weedon et al (2005) extended the investigation into common mutations of the GCK gene and reported that the maternal A allele of SNP rs1799884 was associated with a 64 g increase in birth weight. No effect of fetal genotype was observed in this study, which included more than 2000 mother child pairs (Weedon et al., 2005).

Maternal T2D risk genotypes at rs7903146 of the transcription factor 7 like 2 (TCF7L2) were associated with increased offspring birth weight (Freathy et al., 2007). Fetal genotypes at T2D risk alleles of TCF7L2 (Mook-Kanamori et al., 2009, Freathy et al., 2007, Cauchi et al., 2007b), peroxisome proliferator-activated receptor gamma (PPAR-γ) and potassium inwardly-rectifying channel, subfamily J, member 11 (KCNJ11) genes (Bennett et al., 2008) were not associated with birth weight. A recent study indicated that fetal genotype at CDK5 regulatory subunit associated protein 1-like 1 (CDKAL1) and homeobox, hematopoietically expressed/insulin degrading enzyme (HHEX-IDE) was associated with reduced birth weight, and that carriers of 4 risk alleles were 80 g lighter compared to non-carriers, an effect that was slightly modified by maternal genotype (Freathy et al., 2009). This latter study gives genetic evidence for the fetal insulin hypothesis at the genetic level which supports the epidemiological data.

In a recent GWAS study including 10,000 individuals, 2 SNPs on fetal chromosome 3 were associated with a 112 g reduction in birth weight, equivalent to the effect of maternal smoking of 4 cigarettes a day. One of the SNPs, rs9883204, was close to the adenylate cyclase 5 (ADCY5) gene. The birth weight lowering C allele of this SNP was in (LD) with allele A of SNP rs11708067 ($r^2 = 0.75$), which is associated with a higher risk of T2D and known to alter insulin secretion, further supportive evidence for the fetal insulin hypothesis (Freathy et al., 2010).

1.2.6.2 Genomic imprinting defects and FGR

Certain genes are subject to parent of origin specific silencing in which paternal or maternal alleles have reduced or absent expression (Beechey, 2004). This silencing takes place through epigenetic mechanisms, resulting in differential levels of transcription with gene expression predominantly from one parental allele.

The placenta is a major site for genomic imprinting. It has been hypothesised that in the placenta maternal and paternal genes show a conflict of interest. Paternal genes are designed to maximize the nutrient supply for the fetus, whilst maternal genes have evolved to limit fetal growth firstly to prevent macrosomia and birth related injury and secondly to maintain resources for subsequent pregnancies. Consequently, imprinted genes are expected to play a significant role in fetal growth (Reik et al., 2003).

At chromosome 11p15 is a cluster of genes that are subject to imprinting under an imprinting control region (ICR) located downstream of insulin-like growth factor II (IGF-II) and upstream of H19 genes. IGF-II is growth promoting and expressed exclusively from the unmethylated paternal allele, whilst H19 is growth limiting and expressed exclusively from the unmethylated maternal allele. Loss of methylation of the paternal allele of the ICR was reported in 60% of cases of Silver-Russell syndrome, an extreme syndrome of intrauterine growth retardation and dysmorphic features (Netchine et al., 2007). Gain of methylation of the maternal copy of ICR was associated with another syndrome called Beckwith-Wiedemann syndrome, a syndrome characterized by pre- and post-natal overgrowth. The placenta from FGR pregnancies was found to exhibit

reduced methylation of this chromosomal region compared to the AGA placenta (Bourque et al., 2010). IGF-II knockout in mouse models leads to FGR and small sized placenta (Sibley et al., 2004). However, IGF-II expression at the level of mRNA in FGR placenta showed contradictory results. Lee et al (2010) reported higher levels of IGF-II mRNA in FGR placenta whilst Guo et al (2008) reported lower levels (Guo et al., 2008, Lee et al., 2010).

1.2.6.3 Other candidate gene studies

Multiple genetic pathways known to regulate placentation or related to a risk factor for FGR have been investigated. Genes for growth factors, cytokines, MMPs, thombophilia, and the renin-angiotensin system, represent the bulk of genetic studies in FGR. The lack of reproducibility of these studies is typical of candidate gene studies of complex disorders. In addition, none of the studies was large enough to test for fetal and maternal gene interaction.

It is beyond the scope of this thesis to review these studies, but studies involving EGF will be considered in more detail, as it is one of the candidate genes to be investigated in chapter 3. EGF is a potential candidate gene for FGR and pre-eclampsia. It is located near to a region on chromosome 4 which has been linked to preeclampsia (Harrison et al., 1997) and birth weight (Arya et al., 2006). EGF regulates many aspects of trophoblast function (Maruo et al., 1992). Dissanayake et al (2007) tested 2 SNPs, (c.2566 G>A and c.61 A>G) in EGF in relation to FGR (Dissanayake et al., 2007). These two SNPs are located in exon 14 and the 5' untranslated region of EGF gene (Shahbazi et al., 2002). The maternal G allele of c.61 A>G and the A allele of c.2566 G>A were associated with lower birth weight in healthy babies. Furthermore, the haplotype formed by these risk alleles

was transmitted more frequently from heterozygous parents to FGR offspring. These findings were evident in both Caucasian and Sinhalese women. The EGF gene and polymorphisms and their role in pregnancy will be discussed in chapter 3.

1.3 Summary and hypothesis

Failure of placental development and/or maternal adaptation to pregnancy due to genetic or environmental factors leads to both maternal and fetal complications including preeclampsia and fetal growth restriction, complicating up to 10% of pregnancies.

There is substantial evidence for bidirectional increased risk between T2D and preeclampsia. Both are regarded as genetically complex disorders and share many pathophysiological changes in terms of insulin resistance, dyslipidaemia, and endothelial dysfunction. Low birth weight is widely accepted to be associated with later development of T2D. It is therefore tempting to speculate that some predisposing genetic factors are shared by these conditions. Over the last decade, investigators have focused on the effect of T2D genes on the variability of birth weight. Very limited and underpowered studies have investigated T2D susceptibility genes as a genetic risk factor for preeclampsia and FGR. In view of the discovery of an increasing number of T2D susceptibility loci, the current study will focus on the role of some of these T2D susceptibility loci in preeclampsia and FGR in a case control genetic study (Chapter 2).

Both preeclampsia and FGR share the same placental pathological features leading to placental ischemia, including abnormal spiral artery remodelling, and

shallow trophoblast invasion. EGF is a biologically important growth factor acting locally on the placenta and regulating many aspects of trophoblast functions. Genetic polymorphisms of the EGF gene especially putatively functional variants that can determine EGF expression may represent a genetic risk factor for preeclampsia and fetal growth restriction. The focus of this study is to investigate 2 putatively functional SNPs (rs4444903, and rs2237051) in the EGF gene on the risk of preeclampsia and fetal growth restriction and to describe the molecular genetic aspects of rs4444903 and rs2237051 SNPs on EGF translation, transcription and splicing (Chapter 3).

2 Type 2 diabetes susceptibility genes and the risk of preeclampsia and fetal growth restriction

2.1 Introduction

Type 2 diabetes (T2D) is a metabolic disorder characterized by pancreatic beta cell dysfunction and insulin resistance (Stumvoll et al., 2005). It is estimated to have affected 2.8% of the population worldwide in 2000 and is expected to affect 4.4% (366 million people), by the year 2030 (Cooper-DeHoff, 2007). This is associated with considerable morbidity, mortality, and health-care expenditure.

The higher concordance of T2D in monozygotic twins than in dizygotic twins (70% vs 30%), the higher risk of T2D in a first-degree relative of a patient with T2D (3-fold compared to the general population) and clustering of the disease in families indicated the clear genetic basis for developing T2D. Despite the clear genetic basis of T2D, early genetic linkage and candidate gene analysis validated only a limited number of genetic variants as T2D risk variants. Two of these susceptibility genes (PPAR, and KCNJ11) were already targets for drug treatment of T2D (Table 2-1). In the past three years, GWAS for human T2D from several European populations have confirmed 25 loci to be associated with T2D (Bonnefond et al., 2010).

In the previous section it has been shown that a bidirectional increased risk between T2D and preeclampsia is commonly reported. These two disorders also share the same pathophysiological changes: hyperinsulinaemia, insulin resistance, dyslipidaemia and endothelial dysfunction. Both preeclampsia and

T2D are known risk factors for CAD. Together these observations suggest that preeclampsia and T2D may share the same genetic risk factors.

For the purpose of this study 5 of the most extensively replicated T2D susceptibility genes were selected to test the hypothesis of a shared genetic basis between T2D and preeclampsia and fetal growth restriction. Table 2-1 presents the genetic approach which identified the T2D susceptibility variants, the suggested mechanisms, and the drugs which target the genes selected. More details about each gene and the variants within each gene to be investigated are provided in the following sections.

Table 2-1: T2D genes included in the present study

transcription factor 7 like 2 (TCF7L2), fat mass and obesity associated gene (FTO), peroxisome proliferator-activated receptor gamma (PPAR-γ), Antisense Non-coding RNA in the INK4 Locus (CDKN2B-AS1), potassium inwardly-rectifying channel, subfamily J, member 11 (KCNJ11)

Gene	Genetic approach	Role in T2D	Drug target
TCF7L2	Exploration of linkage peaks	Decreased b-cell function, decreased incretin stimulated insulin secretion, decreased glucose stimulated insulin secretion	
KCNJ11	Candidate gene	Decreased b-cell function, decreased glucose-stimulated insulin secretion	Sulfonylurea drugs
PPARG	Candidate gene	Decreased insulin sensitivity	Thiazolidinediones
FTO	GWAS	Increased risk of obesity, increased body mass index, decreased insulin sensitivity.	
CDKN2B-AS1	GWAS	Decreased b-cell function, decreased glucose-stimulated insulin secretion	

2.2 Candidate genes

2.2.1 Transcription factor 7 - like 2 (TCF7L2)

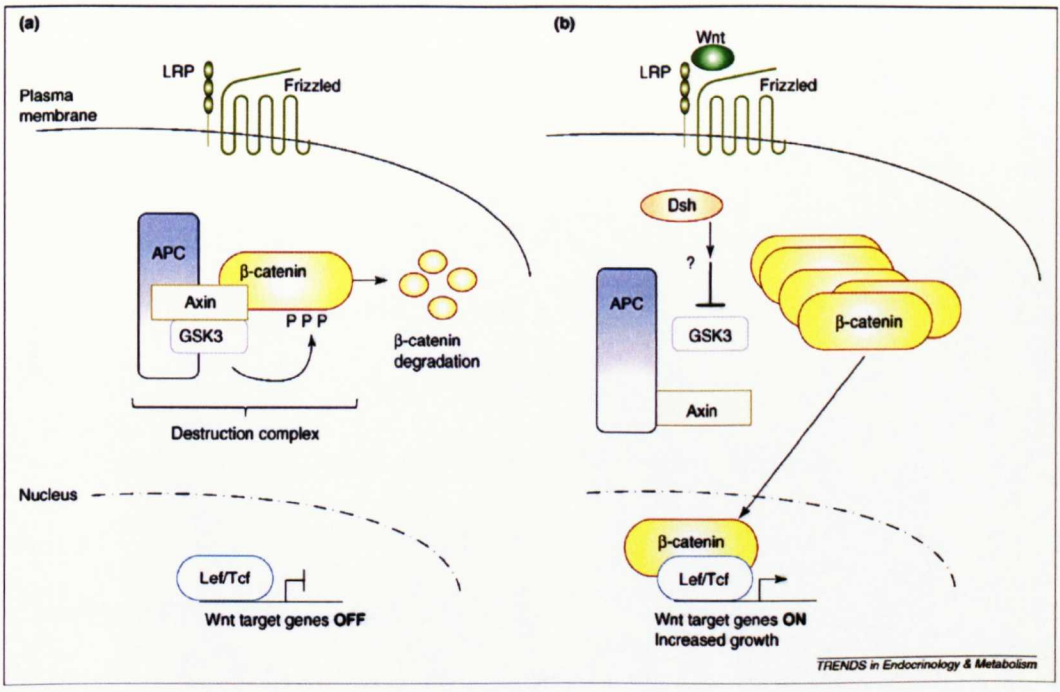
2.2.2 Introduction

Transcription factor 7 - like 2 (TCF7L2) is a transcription factor that belongs to a family of TCF/LEF (T cell-specific factors/ lymphoid enhancer-binding factor) transcription factors. TCF7L2 (also known as TCF-4) interacts with beta (b)-catenin and regulates the WNT (wingless) pathway (Jin and Liu, 2008). Consequently it is involved in growth and differentiation from embryonic life to cancer pathology.

In the absence of WNT ligands, TCF7L2 which is localized mainly in the nucleus acts as a repressor for WNT target genes by recruiting other repressors like Groucho and C terminal binding protein-1(C-TBP-1) to the promoter regions of these genes. When levels of WNT ligands are low, the cytoplasmic b-catenin undergoes proteolytic degradation subsequent to phosphorylation by glycogen synthase kinase 3 (GSK-3), in the presence of adenomatous polyposis coli (APC) and Axin. When WNT ligands are present, they bind to the seven transmembrane domain frizzled receptors and the low density lipoprotein receptor related protein 5 and 6 (LRP5/6) resulting in dislocalization of a molecule called Dishevelled (Dvl) leading to disruption of the Axin-APC-GSK-3 complex. The dephosphorylated b-catenin is then transmitted to the nucleus and complexes with TCF7L2 resulting in the induction of the target genes (Welters and Kulkarni, 2008, Bodnar et al., 2005), (Figure 2-1).

Figure 2-1: Canonical B-catenin pathway

A) In the absence of WNT signalling. (B) Activation of the WNT-signalling pathway. APC: adenomatous polyposis coli, GSK3: glycogen synthase kinase 3, frizzled: cell-surface frizzled receptors, LRP: low-density lipoprotein receptor-related protein, Dsh: dishevelled, Lef: lymphoid enhancer factor, TCF: T-cell factor (adapted from (Welters and Kulkarni, 2008), with permission.



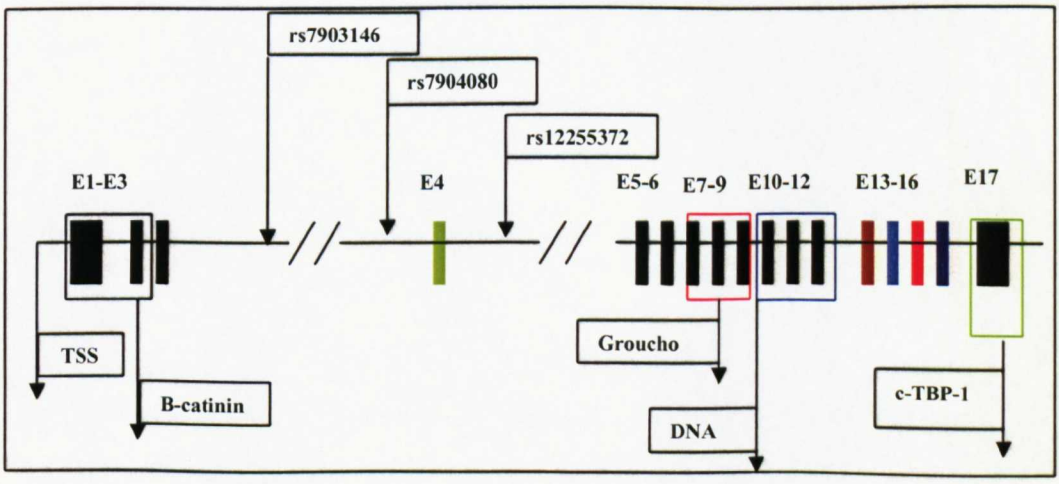
2.2.2.1 Genomic structure

TCF7L2 gene maps to 10q25.3 and comprises 217 Kb. The gene is formed of 17 exons and exhibits a complex pattern of splicing. The protein product comprises many binding domains which bidirectionally regulate WNT signalling pathways (Duval et al., 2000). As demonstrated in Figure 2-2, the b-catenin binding site is encoded by exons 1 and 2 while the c -terminal end of the gene (exon 17) codes the (C-TBP-1) domains. Exons 7-9 produce the binding domain for another co-repressor, called Groucho while exons 10-12 code for the high mobility group (HMG) box DNA binding domain (Osmark et al., 2009), (Figure 2-2) .

Although 3 transcription start sites (TSS) have been identified, due to alternative promoter usage, only one (TSS1) located 536 bp upstream of the translation start site in exon 1, can lead to a transcript with the b-catenin binding domain (encoded by exons 1 and 2). The other two TSS located in exon 1 and intron 1 use a translation start site located in exon 3 (Figure 2-2). Exons 4, and 13-16 are alternatively spliced, giving a high number of splice variants, although only a few variants are expressed in each tissue. Inclusion of exons 14, 15 and 16 results in the introduction of a stop codon and consequently lack of exon 17 from the translation product. The lack of exon 17 leads to truncated protein lacking the co-repressor C-TBP-1 binding domain (Prokunina-Olsson et al., 2009). TCF7L2 has the highest expression in pancreatic islets and lowest expression in skeletal muscle. It is also expressed in gut, liver, peripheral blood cells, visceral and subcutaneous fat (Osmark et al., 2009).

Figure 2-2: Genomic structure of the TCF7L2

TSS: transcription start site 1. E= exon, coloured rectangles indicate the alternative exons (E4, E13-16). Different binding domains in relation to their coding exons are demonstrated by open coloured boxes. The sites of the SNPs to be included in this study are also indicated.



2.2.2.2 Polymorphisms of the TCF7L2

Evidence for the contribution of TCF7L2 to the risk of type 2 diabetes was first provided by a genome-wide linkage analysis in Icelandic pedigrees, in which a linkage peak (10.5 Mb) on chromosome 10 was identified. A fine mapping approach, applying high density microsatellite markers, succeeded in identifying a tetra-nucleotide microsatellite marker (DG10S478) in a reasonably large number of cases and controls of Icelandic descent (1185 cases and 931 controls) (Grant et al., 2006). This marker has six alleles, with alleles 0, 8 and 12 represented in 98% of the population where the 0 allele is protective and non 0 allele (referred to as allele X) is the risk allele for T2D. The same findings were replicated in European American and Danish populations with the same effect size and same level of significance. The relative risk of all 3 population combined was 1.56 (95% CI 1.41-1.73, $P = 4.7 \times 10^{-18}$).

This microsatellite marker resides in the LD block of 92 Kb encompassing intron 3, exon 4 and intron 4 of the TCF7L2 gene on 10q25.3, with no other genes mapped to this region. Sequencing of this region revealed 5 non coding SNPs which were correlated with the marker. Of these SNPs, rs12255372 (r^2=0.95) and rs7903146 (r^2=0.78) showed the same effect on the risk of type 2 diabetes as the original marker with both an additive and a multiplicative mode of inheritance. Population attributable risk (PAR) was 21%, given a MAF of 26% in the study cohort of 3 populations combined (Grant et al., 2006).

Replication studies are typically required in populations of the same and different ethnic groups. This is to confirm the earlier findings regarding the

effect size and also to test for the association with intermediate traits of T2D which helps to understand the exact role of TCF7L2 on the pathology of T2D.

Subsequent studies over the 2 years following the initial study by Grant et al (2006) in European populations as well as populations of non European origin confirmed TCF7L2 as a major T2D susceptibility gene with PAR ranging from 16-30% (Groves et al., 2006, Cauchi et al., 2006). In a case control study selected on the basis of a history of hypo-glycaemic treatment and/or biochemical evidence of diabetes, the T allele of SNP rs7903146 was associated with diabetes in a UK population (Groves et al., 2006). The genotype relative risk of heterozygous TC and homozygous TT carriers was 1.35 (95% CI 1.19 – 1.53, P= 3.1 X 10^{-6}) and 1.90 (95% CI 1.54 –2.33, P= 3.6 X 10^{-6}), respectively.

In a French population the effect of rs12255372 and rs7903146 on the risk of T2D was confirmed. The T allele of the SNP rs7902146 was associated with an average decrease of 2.4 years in the age of onset of T2D (Cauchi et al., 2006). Even in a Finnish population with lower MAF for the risk alleles (17% compared to 26%), both rs12255372 and rs7903146 are associated with T2D (Scott et al., 2006). A meta-analysis of candidate gene studies carried by Cauchi et al (2007) found that the pooled OR for the T allele of the SNP rs7903146 was 1.46 (95%CI 1.42–1.51, P=5.4× 10^{-140}). This study included 29,195 controls and 17,202 diabetic cases from 27 studies from different ethnic groups with no heterogeneity in genotypic distribution (Cauchi et al., 2007a).

In addition to the findings of candidate gene studies, several GWAS have identified the SNP rs7903146 as a diabetes risk allele (WTCCC., 2007, Saxena et al., 2007, Sladek et al., 2007, Scott et al., 2006, Zeggini et al., 2007).

Even in populations of Asian descent in which the T allele of rs7903146 has MAF of 0.03, compared to 0.28 in European, it showed significant associations with T2D (Hayashi et al., 2007). Additional independent SNPs emerged in Asian populations including rs11196218 in intron 4 and rs290478 in intron 7 as risk variants for T2D in the TCF7L2 gene (Ng et al., 2007).

Two studies found no association between TCF7L2 variants and the quantitative traits for diabetes, fasting insulin and glucose levels (Cauchi et al., 2006, Groves et al., 2006). However, Saxena et al (2006) found an association between the T allele of the SNP rs7903146 and reduced insulin secretion (50%) as measured by an oral glucose tolerance test (OGTT) in non diabetic Scandinavian carriers (Saxena et al., 2006).

In relation to BMI, the T allele of rs7903146 was associated with lower BMI both in healthy carriers and T2D subjects. Helgason et al (2007) observed that there are two haplotypes of the TCF7L2 gene. HapB, which carries the T2D risk allele of rs7903146 (T allele) was associated with low BMI. HapA, tagged by the C of the rs7903146 and the T allele of SNP rs7904080 was associated with higher BMI (Helgason et al., 2007). Later studies did not find any association between TCF7L2 variants and obesity measures (Pecioska et al., 2010, Cauchi et al., 2008), and suggested that the lower BMI in carriers of the rs7903146 T allele is related to recruitment criteria, and that carriers of the T allele develop diabetes

through effects on insulin secretion, whilst obese individuals would develop diabetes through insulin resistance.

The association between rs7903146 and other diseases has also been tested. For coronary artery disease the results are controversial and less clear than the results obtained with T2D. In a mixed population from the US the T allele of rs7901346 was not associated with either CAD or with peripheral artery disease (Bielinski et al., 2008). In contrast, in a group of 600 non diabetic Brazilian subjects the T allele was associated with CAD (Sousa et al., 2009).

2.2.2.3 Role of TCF7L2 gene in diabetes

TCF7L2 as an effector molecule in the WNT pathway is expected to mainly alter insulin secretion. Oral glucose intake stimulates the secretion of gastrointestinal hormones such as glucagon like peptide -1 (GLP-1) from the distal part of the small intestine, which leads to an increase in glucose stimulated insulin secretion from the b-cells of the pancreas, known as the Incretin effect. It has been reported that TCF7L2 increases the expression of GLP-1 by binding to its promoter. It has been shown that the Incretin effect is reduced in T2D patients as well as in healthy carriers of the diabetes risk alleles (Lyssenko et al., 2007, Helgason et al., 2007). Depletion of TCF7L2 by siRNA leads to reduction of glucose and GLP-1 stimulated insulin secretion (Shu et al., 2008, Shu et al., 2009). Not only was insulin secretion impaired by depletion of TCF7L2, but the whole process of beta cell proliferation and survival were disturbed (Shu et al., 2009). Interestingly, Exendin-4, a GLP-1 receptor agonist, is now in use for the treatment of T2D (Van Gaal et al., 2008).

TCF7L2 is characterized by a complex form of tissue specific splicing. Being intronic in nature and not correlated with any coding SNPs, SNPs of the TCF7L2 gene reproducibly associated with T2D are expected to alter gene expression through mRNA splicing. Transcripts including exon 4 (located just downstream to SNP rs7903146 and predominantly included in the transcript from pancreatic islets), were positively correlated with HbA1c (Osmark et al., 2009). However, the genotype at rs7903146 showed no effect on the expression of any TCF7L2 transcripts, including exon 4 inclusive transcripts in islets taken from 17 non diabetic individuals (Osmark et al., 2009). In another study, the T allele of rs7903146 and the G allele of rs1225572 were negatively correlated with the transcripts including the C terminal exons in 50 islets derived from non diabetic individuals (Prokunina-Olsson et al., 2009). Inclusion of these alternative exons will introduce a stop codon and eliminate the C-TBP-1 binding domain which is a negative regulator of the WNT pathway.

2.2.2.4 Role of TCF7L2 variants in pregnancy disorders

During pregnancy pancreatic b-cells undergo structural and functional changes in response to the increased insulin requirements. Failure of the b-cell of the pancreas to compensate for the increased requirement for insulin during pregnancy leads to gestational diabetes mellitus (GDM).

Some reports are currently available regarding the role of the TCF7L2 variants and GDM. The T allele of SNP rs12255372 in TCF7L2 was significantly associated with GDM in American Mexican women (OR 2.49, 95% CI 1.17–

5.31, P= 0.018) and remained significant when adjusting for age and BMI (OR 2.62, 95% CI 1.13– 6.11, P= 0.025) (Watanabe et al., 2007). In Swedish women, the T-allele of SNP rs7903146 was associated with an increased risk of gestational diabetes mellitus (OR 1.49, 95% CI 1.28–1.75, $P=4.9\times10^{-7}$) (Shaat et al., 2007). This was also reported in Danish women (Lauenborg et al., 2009).

Freathy et al (2007) demonstrated, in large study of a Finnish birth cohort, that the T allele of rs7903146 of both fetal and maternal origin is associated with birth weight. As the fetal and maternal genotypes are 50% correlated, adjustment to fetal genotype showed that the increase in birth weight is due to maternal genotype (31 g/ per risk allele, 95% CI 9–48 g, corrected P= 0.003). The fact that the sample size was large and the effect on birth weight was supported by the finding that the carriers of the risk allele had reduced insulin secretion and high fasting glucose makes this association robust (Freathy et al., 2007).

Fetal SNP genotype did not show any effect on intrauterine weight gain or 2 years postnatal weight gain in Dutch subjects either in the general population or in a small for gestational age cohort (Mook-Kanamori et al., 2009). A French study found no association between the T allele of rs7903146 and smallness for gestational age (Cauchi et al., 2007b). The lack of effect of fetal genotype at this locus on birth weight was explained by the fact that even if insulin secretion was reduced in a carrier of the risk allele it would be of a modest nature and of too short duration to lead to SGA. Although these studies excluded an effect of fetal genotype at this locus on birth weight or risk of SGA, they did not study the effect of maternal genotype or the effect of other variants of this gene or interaction with other T2D susceptibility genes on birth weight.

TCF7L2 has not been tested in preeclampsia and it merits investigation both as a diabetes gene and also as a part of pathway that is reported to be expressed in the placenta and to regulate the invasive function of the trophoblast (Pollheimer et al., 2006).

2.2.3 Fat Mass- And Obesity-Associated Gene (FTO)

2.2.3.1 Introduction

FTO, fat mass and obesity-associated gene, is ubiquitously expressed in human tissues with high expression in the brain, especially in the region of the hypothalamus and its nuclei that regulate energy expenditure and food intake. It is also highly expressed in both visceral and subcutaneous adipose tissue (Frayling et al., 2007, Boissel et al., 2009). The gene encodes α-2-oxoglutarate–dependent nucleic acid demethylase, responsible for nucleic acid demethylation (Gerken et al., 2007), suggesting that it may operate through epigenetic mechanisms.

2.2.3.2 Genomic structure and polymorphisms of FTO gene

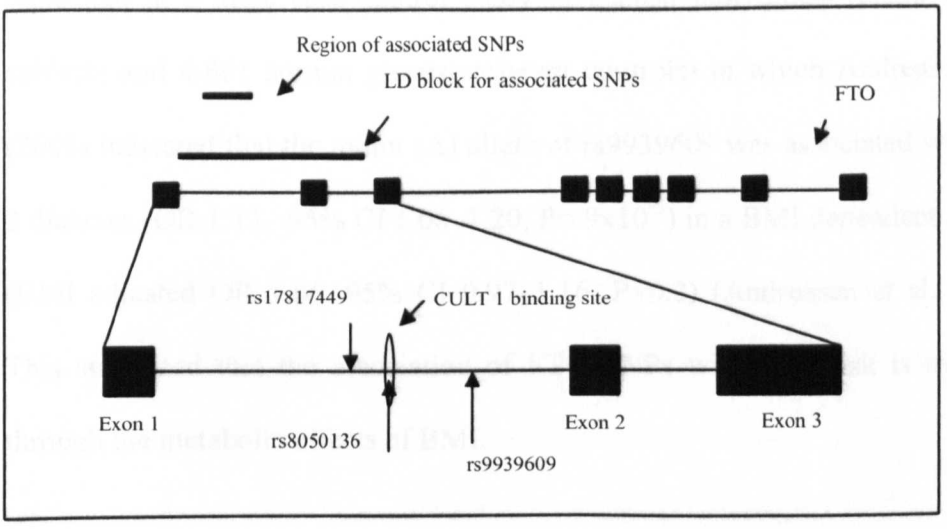
The FTO gene is a large gene formed of 417 Kb mapped to chromosome 16q. It comprises 9 coding exons generating a 1518 bp transcript and the protein product of 505 AA.

Part of the large intron 1 (105kb) along with exon 2 (78bp) and part of the intron 2 (15kb) form a 47 kb LD block in which all identified T2D and obesity risk variants of the FTO gene are located (Figure 2-3). This LD block carries a

cluster of 40 SNPs that are highly correlated (linkage disequilibrium $r^2 \geq 0.80$) in Caucasian populations. The fact that linkage disequilibrium decays markedly beyond this region indicates that the authentic and /or causal allele lie in this region.

Figure 2-3: Organization of the FTO gene

FTO gene id formed of 9 exons (Black boxes). A 47 Kb LD block carrying the T2D and obesity SNPs is shown spanning intron 1, exon 2 and most of intron 2. The relative location of the most commonly studied SNPs and a CULT1 transcription factor binding site are shown. Adapted from (Stratigopoulos et al., 2008).



2.2.3.3 Role of FTO in Diabetes, obesity and metabolic syndrome

Obesity is a major risk factor for diabetes, cardiovascular disease, metabolic syndrome, hypertension, and preeclampsia. A GWAS for T2D identified 2 common variants (rs9939609, rs8050136) in the FTO gene that predispose to diabetes through an effect on BMI (Scott et al., 2007, WTCCC., 2007). These effects on the risk of T2D, obesity, and their related traits were evident in children and adults of both genders and in different populations.

The Wellcome Trust Case Control Consortium (WTCCC), that included 1924 T2D subjects and 2938 population controls (British individuals) showed that the A allele of SNP rs9939609 in FTO was related to an increased risk of T2D (OR 1.27, 95% CI 1.16-1.37, $P=5\times10^{-8}$) (WTCCC., 2007). This association was replicated by genotyping of a further 3757 T2D cases and 5346 controls (OR 1.15, 95% CI 1.09-1.23, $P=9\times10^{-8}$). The association was eliminated after adjusting for BMI (OR1.03, 95% CI 0.96-1.10, P=0.44) (Zeggini et al., 2007). The interaction between BMI and FTO genotype on the risk of T2D was also replicated in a large case control study in Danish individuals (3,856 diabetic subjects and 4,861 normal glucose-tolerant controls) in which Andreasen et al (2008) indicated that the minor (A) allele of rs9939609 was associated with type 2 diabetes (OR 1.13, 95% CI 1.06–1.20, $P= 9x10^{-5}$) in a BMI dependent manner (BMI adjusted OR 1.06, 95% CI 0.97–1.16, P=0.2) (Andreasen et al., 2008). This suggested that the association of FTO SNPs with T2D risk is mediated through the metabolic effects of BMI.

Many studies have confirmed an association of FTO with obesity. Dina et al extensively studied FTO variants in relation to obesity and recorded a population attributable risk (PAR) of 22% using a multinational study comprising 2,900 obese individuals and 5,100 controls of European ancestry (French, German, Swiss, and Swedish). The C allele of SNP rs1421085 and G allele of rs17817449 gave the strongest signal for this association. Furthermore, these alleles were transmitted more frequently from heterozygous parents to obese offspring in French and Swedish families, confirming them as obesity risk variants (Dina et al., 2007). Frayling et al studied the association of FTO gene variation with increased BMI and the risk of being overweight and obese in 19,424 white

European adults from seven general population-based studies (mean age 28 to 74 years, mean BMI 22.7-27.2 kg/m^2) (Frayling et al., 2007). The OR per-A allele of SNP rs9939609 was 1.31 (95% CI 1.23-1.39, P=6x10^{-16}) for obesity in the adult general population and 1.18 (95% CI 1.13-1.24, P=1x10^{-12}) for being overweight. The same effect was manifested in 10,172 white European children, mean age 7 to 14 years, mean BMI 16.1 to 19.2 kg/m^2).

Variants of FTO have been also associated with the quantitative traits of metabolic syndrome (MS), T2D and obesity. In 17000 individuals, FTO genetic variants were associated with characteristics of metabolic syndrome (high fasting insulin, glucose and triglycerides, and low HDL) (OR 1.17, 95% CI 1.2-1.25, P=3x10^{-6}) (Freathy et al., 2008). This association with metabolic traits was abolished after adjusting for BMI, indicating that the increase in the body fat in carriers of the risk allele may impose these metabolic changes. Similarly, Do et al reported an association between rs1421085 and rs17817449 and elevated fasting insulin levels and insulin sensitivity (diabetes traits) in a BMI dependent manner (Do et al., 2008). The A allele of rs9939609 was also significantly associated with higher plasma insulin levels (P=0.05), higher insulin resistance index (homeostasis model assessment) (P=0.02), and higher systolic blood pressure (P=0.003) in a French population, but these associations disappeared after adjustment for BMI (Legry et al., 2009). In both of these large genetic studies an association between rs17817449, rs1421085 and rs9939609 and obesity quantitative traits was also evident (P= 0.00008 for BMI, P=0.0002 for waist circumference, and P=0.006 for body weight) (Legry et al., 2009, Do et al., 2008, Andreasen et al., 2008). Andreasen et al demonstrated gene environment

interaction, such that higher insulin sensitivity attenuates and low physical activity accentuates the effect at rs9939609 on obesity (Andreasen et al., 2008).

2.2.3.4 FTO variants and CAD risk

In relation to hypertension, Pausova et al showed that rs9939609 was associated with higher blood pressure independent of BMI and this was evident in both an adolescent sample and an independent adult sample (4.4 mm Hg/allele, P=0.0024), indicating that individuals at risk for hypertension may be recognized early during development of the disease when preventive measures may still be effective (Pausova et al., 2009). The authors suggested that the association of this variant with higher blood pressure is related to an increase in sympathetic vasomotor tone as it was more evident when tested under mental stress (Pausova et al., 2009). This is consistent with the fact that FTO is expressed at higher levels in the hypothalamic centres known to regulate blood pressure. FTO variants were also associated with systolic blood pressure in a French population (P=0.003, BMI dependent) (Legry et al., 2009) and with systolic and diastolic blood pressure in Finnish women (Lappalainen et al., 2010). In this latter study, it was also possible to show the association of this variant with increased risk of CAD in men (OR 2.09 95% CI 1.17-3.73, P=0.013). The association was independent of BMI or C reactive protein levels, but was partially explained by increased RANTES (interleukin 8 related cytokine) in these men who either displayed impaired glucose tolerance or were diabetic (Lappalainen et al., 2010). This confirmed the earlier findings by Doney et al who introduced for the first time the association of FTO variants with an increase in risk of CAD in diabetic Scottish individuals, (OR 2.0, 95% CI, 1.18 -3.45, P=0.011) (Doney et al., 2009).

The association was actually attenuated after adjustment for BMI but remained significant. In both of these studies the FTO risk variant was associated with an atherogenic lipid profile (lower HDL, high triglyceride levels), as well as with inflammatory markers (C-reactive protein and RANTES). Besides the metabolic changes, the atherogenic lipid profile, and inflammatory changes associated with FTO risk variants, FTO is currently reported to be associated with altered nucleic acid methylation, an epigenetic change that appears to be enhanced in CAD (Sharma et al., 2008, Han et al., 2010).

The association of FTO variants and T2D, obesity, and metabolic syndrome was also reported in other ethnic groups like Chinese (Han et al., 2010) and mixed Asian population inhabitants of south Canada (Al-Attar et al., 2008).

2.2.3.5 Functional effects of the FTO variants

In view of the large number of studies that relate FTO variants to obesity, T2D, metabolic syndrome and CAD, it is important to find out how variations in the FTO gene contribute to disease development. Since the substrate for the FTO encoded enzyme (α-2-oxoglutarate–dependent nucleic acid demethylase) is not yet identified, the exact function of FTO is not fully understood. Most studies have been directed towards identification of the function of FTO itself and few have studied the role of the risk variants.

In humans, it appears that hyperphagia and food preference toward high energy diet are the causes of obesity (Cecil et al., 2008, Speakman et al., 2008). Loss of the mouse FTO gene product resulted in reduced postnatal weight gain, reduced

adiposity, increased energy expenditure due to sympathetic stimulation, and hypophagia (Church et al., 2009, Fischer et al., 2009). These data suggested that variants in the human FTO gene which lead to an altered expression (up-regulation) might render individuals more susceptible to obesity. In agreement with this hypothesis FTO showed higher expression in the subcutaneous and visceral fat from obese compared to control subjects (Zabena et al., 2009, Wahlen et al., 2008). Although no impact of the genotype at rs8050136 and FTO expression was reported in a sample of 306 women (with BMI 18–53 kg/m^2), the FTO variant rs8050136 was associated with altered rates of lipolysis both *in vivo* (the T allele was associated with a 30% increase of serum glycerol) and in isolated adipocytes (Wahlen et al., 2008). Of note, the binding sites for CUTL1, a transcriptional regulator, overlaps with rs17817449 and rs8050136; blocking these sites was associated with altered FTO expression in fibroblast (Stratigopoulos et al., 2008).

In summary, altered FTO expression leading to increased food intake, reduced energy expenditure, altered lipolysis/lipogenesis, increased peripheral insulin resistance, an increased systemic inflammatory state, and altered epigenetic effects, are suggested mechanisms by which the association between FTO genetic variants and disease status can be explained.

2.2.3.6 Role of FTO in pregnancy

The consistent association between the FTO gene and obesity and BMI and the known correlation between adult BMI and birth weight render the FTO gene of interest for birth weight association. The variant rs9939609 was not associated

with birth weight in a small genetic study involving 214 Belgian newborns, but was associated with postnatal weight gain at 2 weeks post partum (Lopez-Bermejo et al., 2008). This was in agreement with an animal study showing that inactivation of FTO was associated with normal birth weight but with reduced weight gain postnatally. Bassols et al reported that expression of FTO in human placental tissue correlated with birth weight (Bassols et al., 2010).

There is no evidence that FTO variants alter the risk of GDM (Lauenborg et al., 2009) which is expected as FTO increases the risk of diabetes by increasing insulin resistance rather than altering the b-cell function which is the main factor in GDM. No reported studies have tested FTO or its variants in relation to preeclampsia risk although it is a strong candidate gene, as it is associated with BMI, T2D, CAD, hypertension, and a systemic inflammatory state, all of which are potential risk factors for preeclampsia.

2.2.4 Peroxisome proliferator-activated receptors-gamma (PPAR-γ)

2.2.4.1 Introduction

Peroxisome proliferator-activated receptors (PPARs) are members of the nuclear hormone receptor superfamily. This family of transcriptional regulators includes steroid hormone, thyroid hormone and Orphan receptors such as PPARs. PPARs represent group C in subfamily 1 of the nuclear hormone receptor superfamily (NR1C). This large subfamily also comprises for instance, All-trans retinoic acid receptors (NR1A), the retinoic acid receptors, (NR1B), and the vitamin D3 receptor, (NR1I). There are three PPAR isotypes, namely PPAR-alpha (NR1C1), PPAR-beta (NR1C2; also called PPAR-delta), and PPAR-gamma (NR1C3).

PPARs are recognized in different species including humans (Escher and Wahli, 2000).

There are different functional domains in the PPAR proteins. The N-terminal A/B domain is a ligand-independent transactivation domain called activation function-1 (AF-1), the highly conserved C domain forms the DNA-binding domain (DBD), and the D domain links the DBD to the c-terminal end of the receptor. The c-terminal end is the ligand-binding domain (LBD), containing a ligand-dependent transactivation function, termed activation function-2 (AF-2). Although the DBD and LBD show 60-80% sequence similarity between different isotypes, they still maintain a certain level of variation for both tissue and ligand specificity (Escher and Wahli, 2000).

PPARs are ligand-activated transcription factors and form heterodimers with retinoic X receptors (RXR) to regulate the transcription of various genes after binding to PPAR response elements (PPREs). The PPRE consists of 2 simple repeats of AGGTCA separated by a single nucleotide. PPARs can be activated by natural ligands, such as prostaglandins (PGs), fatty acids and their derivatives, as well as by synthetic ligands like thiazolidinedione drug derivatives (troglitazone, pioglitazone, and rosiglitazone). In the absence of ligands, PPARs induce a repressor function on the target genes by binding to the promoter region of these genes, interacting with other co-repressors. On ligand binding, the PPAR/RXR heterodimer associates with cofactors containing histone acetyl-transferase activity, modifying nucleosome structure and contacting general transcription factors with induction of the target genes (Bragt and Popeijus, 2008).

The effect of PPARs on the target promoters depends on the phosphorylation status of the receptors (that alter the affinity of the receptor for DNA, RXR, and ligand), type and availability of ligands, DNA allosteric constraints, promoter context and the cell-type specific combination of transcriptional co-regulators that bind to the PPARs/RXR heterodimer (Tan et al., 2005).

PPAR isotypes are differentially expressed in tissue. PPAR-beta is ubiquitously expressed, PPAR-alpha is expressed in tissues with a high metabolic rate like liver, heart, kidney, and skeletal muscle, while PPAR-gamma (PPAR-γ) is mainly expressed in adipose tissue and to a lesser extent in skeletal muscle, colon and placenta (Escher and Wahli, 2000).

PPAR-γ is an important adipogenic regulator and modulator of glucose metabolism. PPAR-γ stimulates adipocyte differentiation and leads to the formation of an increased number of small adipocytes. These small adipocytes are able to take up free fatty acids more easily and therefore decrease the flux of free fatty acids to the liver and muscle. The lowered flux of free fatty acids into skeletal muscle leads to better uptake and use of glucose as a source of energy. This adipogenic effect is mediated through activation of specific genes like CD36, a gene belonging to the class B scavenger receptor superfamily, and lipoprotein lipase. PPAR-γ activation increases skeletal and hepatic glucose uptake and reduces hepatic glucose production (Honka et al., 2009). PPAR-γ also has anti-inflammatory effects through reducing the production of proinflammatory cytokines after forming complexes with Nf-KB.

Thiazolidinediones, potent and relatively specific ligand activators of PPAR-γ currently used for the treatment of type 2 diabetes, lead to reduced fasting insulin and glucose concentrations, and improved whole body insulin sensitivity (Bragt and Popeijus, 2008).

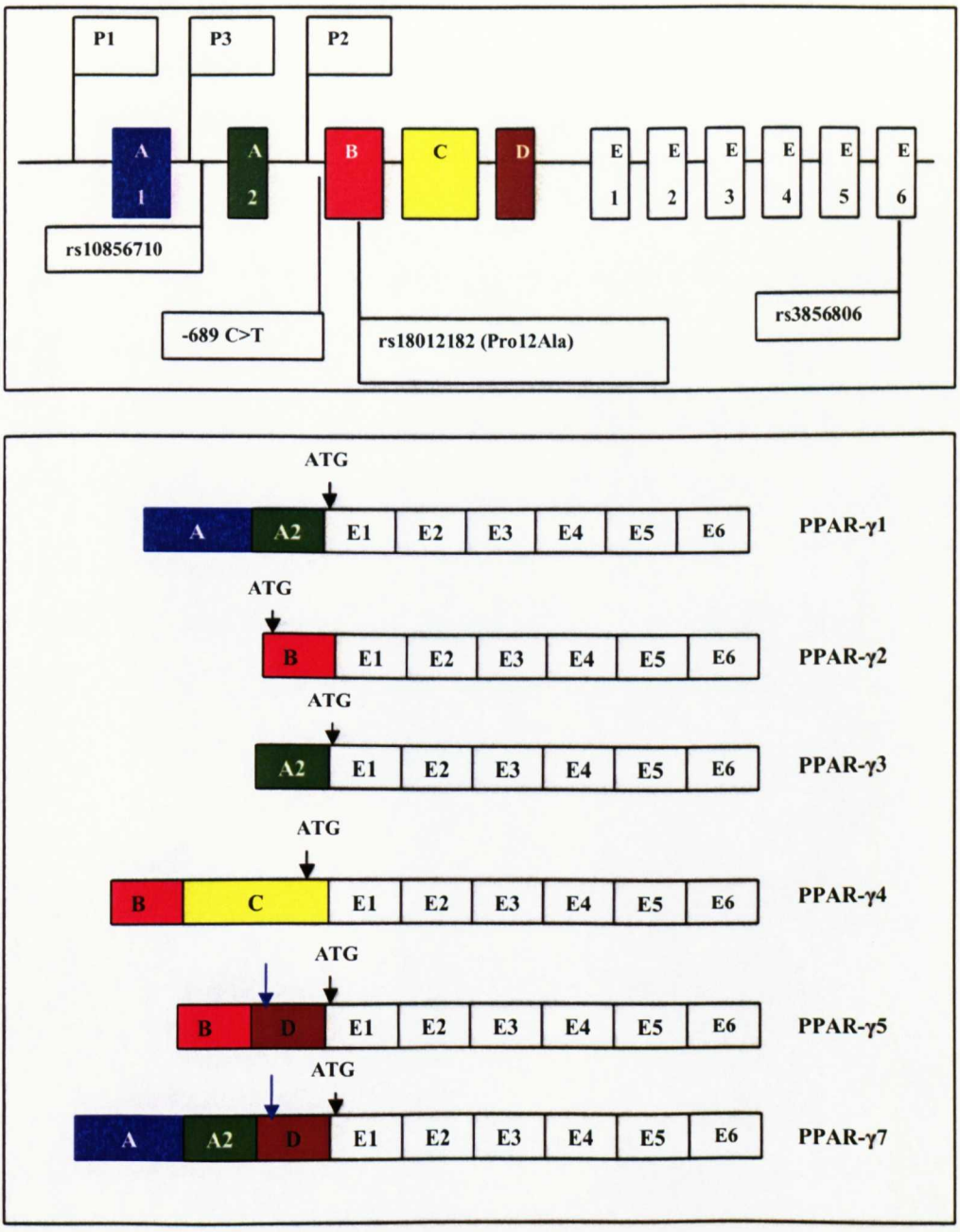
2.2.4.2 Genomic structure of PPAR-γ

Each of the PPAR isotypes is encoded by a different gene; for example, PPAR – alpha is on human chromosome 22, while PPAR-beta maps to chromosomes 6. The PPAR-γ gene is located on chromosome 3, at position 3p25 spanning around 125Kb (Beamer et al., 1997). There are 6 transcripts of PPAR-γ which have been identified in humans, formed due to different promoter usage and alternative splicing. These are PPAR-γ1, PPAR-γ2, PPAR-γ3, PPAR-γ4, PPAR-γ5 and PPAR-γ7. As demonstrated in Figure 2-4, six coding exons (E1-E6) are recognized in which exon 1 encodes for the A/B domain, exons 2 and 3 encode for DBD, while exon 5 and 6 encodes for LBD. The 5'-terminal region of the transcript is the determinant of the PPAR-γ isoform. Five alternative exons have been identified in the 5'-terminal region, referred to as exons A1, A2, B, C, and D. They are alternatively spliced with exons 1–6 to generate the 6 transcripts of PPAR-γ as shown in Figure 2-4. PPAR-γ1, PPARγ-3, PPAR-γ5 and PPAR-γ7 are different transcripts, but encode the same protein (475 AA and 54.5 kDa) as they share the same translation start site in exon 1. PPAR-γ1 consists of exon A1 and A2 spliced together with exons 1–6. The mRNA for PPAR-γ2 consists of exon B and exons 1–6. The third isoform, PPAR-γ3, identified in humans, consists of only exon A2 in its c-terminal region and exons 1–6. There is an ATG in exon B, however in PPAR-γ5 and PPAR-γ7, this site is inactivated by

stop codons present in exon D. Thus, exon B contributes 28 n-terminal AA in PPAR-γ2 only (505 AA, 57.6 kDa). A translation initiating ATG is also present in exon C resulting in an additional eight AA at the n-terminus of PPAR-γ4 (483 AA, 55.3 KD). So the six transcripts encode a total of 3 different PPAR-γ protein isoforms due to different promoter usage and alternative splicing (Elbrecht et al., 1996, Chen et al., 2006). Although alternative splicing is more commonly known to provide different proteins from the same DNA, the case here is that different transcripts with different lengths of the 5'UTR regulate the efficiency of protein translation from different PPAR-γ transcripts (McClelland et al., 2009). Tissue specificity is reported to govern which transcript is predominantly expressed in a given tissue; PPAR-γ2 and PPAR-γ3 are predominantly expressed in adipose tissue whereas PPAR-γ1 expression is ubiquitous. Expression of isoform-specific transcripts may also be regulated by PPAR-γ ligands. For example, in human macrophages, treatment with 15-deoxy-12, 14-prostaglandin J_2 (15d-PGJ2) resulted in a 22% induction of PPAR-γ5/-γ7 transcript and a 24% repression in PPAR-γ4 transcript. Troglitazone was effective at repressing PPAR-γ4 levels by 23% but it did not significantly induce the PPAR-γ5/γ7 transcript. Interestingly, both 15d-PGJ and troglitazone treatment significantly inhibited the expression of PPAR-γ1 and PPAR-γ2 by 30 and 70% respectively compared to untreated controls (Chen et al., 2006).

Figure 2-4: Genomic structure of PPAR-γ

Upper panel showing the coding exons (E1-E6), the alternative 5'-terminal exons (A1, A2, B, C, and D), alternative promoters (P1-P3), and the site of SNPs to be investigated in this study. Lower panel, showing the different PPAR-γ transcripts and the translation initiation site in each transcript (black arrow). The site of the stop codon in exon D is also shown (blue arrows).



2.2.4.3 Polymorphisms of the PPAR-γ

Multiple rare mutations have been identified in the PPAR-γ gene that provided evidence for the central role of this gene in glucose and fat metabolism. V290M, R425C, P467L, and F388L (located in the LBD) have been associated with PPAR ligand resistance (PLR) syndrome. This syndrome is characterized by partial lipodystrophy (loss of subcutaneous fat from certain parts of the body and preserved in other parts), early-onset severe insulin resistance (IR), type 2 diabetes, dyslipidaemia (high triglycerides, low HDL cholesterol), early-onset hypertension and hepatic steatosis. These mutations render the receptor less active. Also a frame-shift mutation that introduces a premature stop codon (186 stop located in the DBD) is described in individuals with severe insulin resistance. P115Q, a specific mutation of the PPAR-γ 2 (located on exon B), was described in 4 German individuals with morbid obesity and type 2 diabetes. This mutation was associated with defective phosphorylation of the receptor at serine 114 and renders the receptor more active with better adipocyte differentiation and lipid accumulation reviewed in (Meirhaeghe and Amouyel, 2004).

Two common polymorphisms were studied extensively based on this prior knowledge of the biological and metabolic functions of the PPAR-γ2: Pro12Ala (rs1801282, C>G) in exon B and C1431T (rs3856806) in exon 6 (Figure 2-4).

Pro12Ala is located in the specific exon B of the PPAR-γ2 transcript and results from CCG>GCG with substitution of alanine for proline in codon 12. The frequency of the Ala allele varies across ethnic populations (1-3% in Africans and Asians and as high as 12% in Caucasians), (Yen et al., 1997).

Deeb et al (1998) were the first to report an association between T2D and Pro12Ala; they showed that the Ala (G) allele is protective for T2D in a small group of Japanese Americans (Deeb et al., 1998). This study was followed by many studies with contradictory findings regarding the protective effect of the Ala allele on the risk of T2D in different ethnic groups, including Danish (Ek et al., 2001), French (Meirhaeghe et al., 2000) and Chinese (Fu et al., 2001).

The low MAF of the Ala allele, the small effect size, small sample size, the effect of population stratification, and gene environment interaction might be the causes of these contradictory results. Adequately powered studies, family based studies, and the more recent genome wide association studies strongly confirmed the Ala allele as a protective allele for T2D.

In a case control study comprising 1000 cases and 1000 controls from the UK population, the Ala allele was found to be protective against T2D (Pro allele OR was 1.4, 95% CI 1.08–1.99, P= 0.01) (Zeggini et al., 2005), and similar results were obtained in a French population (Pro allele OR was 1.37, 95% CI 1.02–1.87, P= 0.04) (Ghoussaini et al., 2005). It was possible to show the interaction between obesity and genotype in relation to type T2D, as restricting the analysis to obese subjects gave a higher OR (Pro allele OR was 1.8, 95% CI 1.5–3.13, P= 0.03) (Ghoussaini et al., 2005). Mori et al also observed the same effects in a Japanese population (Mori et al., 2001).

By using Scandinavian family-based studies as well as conducting a meta-analysis including over 3000 individuals, Altshuler et al (2005) reported that the

Ala allele was indeed protective against T2D and the Pro allele was a risk factor. The PAR was estimated as 0.25 (Altshuler, 2000).

In a study screening 114 SNPs in 76 genes previously reported to be associated with T2D in 786 T2D cases and 617 control subjects with normal glucose tolerance from Finland, Pro12Ala was one of 12 SNPs that replicated as a T2D variant (OR: 1.30, 95% CI : 1.10–1.53, P= 0.009) (Willer et al., 2007). Many of these studies demonstrated that the Ala allele reduced the risk for T2D through improved insulin sensitivity (see below). Further supporting the role of this variant in T2D, the Ala allele was less frequent in those who develop T2D over a 6 year follow up in two large observational and intervention studies carried out in Finland (Lyssenko et al., 2005).

In relation to BMI contradictory results have been reported. Some studies reported the Ala to be associated with a low BMI (Deeb et al., 1998, Doney et al., 2002), whilst others reported a high BMI in Ala allele carriers (Meirhaeghe et al., 2000). Masud et al (2003) showed that even though Ala allele carriers had higher BMI, only approximately 1% of the total variance of BMI can be explained by this polymorphism (Masud and Ye, 2003). In the same study by Masud et al, a meta-analysis using data from 30 independent studies with a total number of 19136 subjects, showed that the Ala allele was associated with a higher BMI only in obese individuals (BMI>27kg/m^2). It has been suggested that homozygosity for the Ala allele is associated with higher BMI in obese subjects, but within the non obese subjects Ala allele homozygotes have a lower BMI than Pro allele carriers (Ek et al., 1999). This suggested that there are modifiers,

possibly environmental or genetic interaction, that alter the effect of this variant on obesity.

The impact of Pro12Ala on obesity is indeed modified by other SNPs in the same gene. As shown by Doney et al (2002), the Ala allele was associated with low BMI while the T allele of the rs3856806 (C1431T) SNP was associated with an opposing effect(Doney et al., 2002). Another two SNPs were also tested in the alternative promoters of PPAR-γ (Figure 2-4) in which the G allele of P3 SNP (rs10865710, C>G) and the T allele of the P2 SNP (-689 C>T) were associated with higher body weight in a French population. Haplotype analysis showed that the haplotype formed of the minor alleles of P3-P2-and Pro12Ala (G-T-Ala) was associated with higher body weight compared with the haplotype formed of the major alleles. Interestingly SNPs in P2 and P3 were not associated with T2D, glucose or insulin variables (Meirhaeghe et al., 2005). A synonymous SNP in exon 6 of the PPAR-γ gene (rs3856806; C1431T) was associated with type 2 diabetes under a dominant model (OR 1.26, 95% CI 1.05–1.51, P= 0.013) but was no longer significant after adjustment for the Pro12Ala variant (OR 1.11, 95% CI 0.89–1.38, P = 0.064), (Willer et al., 2007).

Both Pro12Ala and C1431T (LD, r^2= 0.7) were found to be associated with CAD. Ridker et al (2003) reported the Ala allele to be protective against CAD in US white men (OR: 0.76, 95% CI 0.59–0.99, P= 0.034), in a study which included 523 cases and 2092 controls (Ridker et al., 2003). The T allele of C1431T SNP was also associated with CAD in an Australian white population (Wang et al., 1999). In both studies the allele attributable risk of CAD was independent of diabetes, body mass index, waist-to-hip ratio, or lipid profiles.

2.2.4.4 Mechanism of action of PPAR-γ variants in T2D

Deeb et al (1998) reported that the Ala12 isoform from the PPAR-γ2 transcript bound less effectively to the PPRE and activated the PPRE-driven reporter gene in HepG2 cells less effectively than the Pro isoform in the presence of increasing concentrations of troglitazone (Deeb et al., 1998). This was replicated in two other cell lines, CV-1 (monocyte) and 3T3-L1 (adipocyte lineage) in both the presence and absence of PPAR-γ2 ligands. Both the Acyl-CoA oxidase and the lipoprotein lipase promoters, previously shown to contain a functional PPRE, were also less efficiently transactivated by the PPAR-γ2 Ala isoform than the Pro isoform in the absence of added ligand (Masugi et al., 2000).

In vivo, insulin sensitivity, as determined by a euglycaemic–hyper-insulinaemic glucose clamp, was improved in bearers of the Ala allele. Moreover, the insulin sensitivity of glucose disposal and insulin sensitivity of lipolysis were greater in the subjects carrying the Ala allele (Stumvoll et al., 2001).

The fact that the less active Ala allele is protective against T2D and associated with better insulin sensitivity is puzzling considering the use of PPAR-γ agonists for treatment of T2D. Actually, endogenous ligand free PPAR-γ (known to be a transcriptional repressor) or PPAR-γ along with its natural ligand might primarily serve to hamper insulin action and induce insulin resistance. Therefore introducing the less active Ala variant or the use of thiazolidinediones, (most of which are both partial agonists and partial antagonists) would restore insulin sensitivity. Thus, reduced transcriptional activity of the Ala isoform in adipose

tissue might enhance insulin's capacity to suppress lipolysis, leading to a decrease in free fatty acid release into the circulation, and forcing the muscle to use more glucose. In support of this, it has been shown that in lean glucose tolerant subjects, there is a decrease in lipid oxidation and increase in carbohydrate oxidation during insulin stimulation in carriers of the Ala allele (Stumvoll et al., 2001).

2.2.4.5 The role of PPAR-γ in pregnancy

PPAR-γ is expressed predominantly in invasive trophoblast of the 1st and 2nd trimester (Schaiff et al., 2000, Waite et al., 2000, Rodie et al., 2005) and in both CT trophoblast and ST trophoblast of the 3rd trimester placenta. Although PPAR-γ is expressed in invasive trophoblast Tarrade et al showed that both natural and synthetic PPAR-γ ligands hampered the invasive function of the primary extravillous trophoblast on matrigel, an effect which is reversed by the use of PPAR or RXR antagonists (Tarrade et al., 2001b). This might indicate that PPAR-γ is required in the quiescent state during which PPAR-γ is known to repress target genes by interacting with promoter-bound co-repressor complexes (this is in keeping with the proposed effect on insulin sensitivity discussed in the previous section).

Augmentation of trophoblast differentiation has important implications for placental function as ST trophoblast is more resistant to hypoxic injury, and carries out important secretory and transport functions in the placenta. Tarrade et al (2001) showed that ligands for PPAR-γ and RXR stimulate differentiation of the ST trophoblast and secretion of human chorionic gonadotrophin (hCG), and

hPL. The authors also demonstrated that the promoter region of the hCG beta gene carries a PPRE and can bind to recombinant RXR and PPAR-γ heterodimers (Tarrade et al., 2001a, Tarrade et al., 2001b). It has been shown that troglitazone stimulates differentiation while 15DPGJ2 inhibits differentiation and increases apoptosis of the trophoblast (Schaiff et al., 2000). This is of great importance as diseases known to be associated with enhanced apoptosis of the trophoblast might benefit from inhibition of 15DPGJ2 production and /or the use of troglitazone (Tarrade et al., 2001a).

The major alterations of glucose and lipid metabolism during pregnancy to meet the increased fetal requirements for glucose indicate that up-regulation of the PPAR-γ system is typically required. Waite et al showed that maternal serum from pregnant women exhibited unidentified factors which stimulate PPAR-γ activity in the Jeg-3 cell line transfected by a luciferase reporter gene under the effect of a PPRE; these factors are absent from the serum of non-pregnant women (Waite et al., 2005). This was interpreted as maternal adaptation for increased glucose metabolism during pregnancy. Later studies by the same group showed that the suspected PPAR-γ activators (probably eicosanoid or fatty acids derivatives as they are resistant to heat coagulation and were extracted from the organic phase of the serum of pregnant women) were expressed at lower concentrations in women with preeclampsia compared to normal pregnancy several weeks before the clinical diagnosis of preeclampsia (Waite et al., 2005). In support of these findings it is well known that the lipid profile of pregnant women is altered markedly during pregnancy with a remarkable increase in prostacyclin (vasodilator) and decrease in prostaglandins (vasoconstrictor) which is reversed in preeclampsia. Changes of free fatty acids and triglyceride levels

and shift in the lipid profile towards oxidized forms are reported in preeclampsia, and may lead to changes in the ligand profile of PPAR-γ as fatty acids are typical ligands for this receptor. Furthermore, PPAR-γ natural (15dPGJ2) and synthetic ligands (troglitazone) were shown to reduce the release of PGE and prostaglandin F2 alpha from the placental and amniotic tissues, PGE and prostaglandin F2 alpha are markers of oxidative stress which are increased in women with preeclampsia (Lappas et al., 2006).

However Rodie et al found no variation in the expression of PPAR-γ in PE, IUGR, or normal placenta by immunohistochemistry and q-PCR (Rodie et al., 2005). It appears that it is the activity of PPAR-γ rather than the level of expression which influences the role in pregnancy related conditions. This activity is governed by the availability and the nature of the ligands for PPAR-γ. These ligands might originate from the maternal circulation, placenta, fetal membranes, amniotic fluid, or decidua.

As PPAR-γ regulates multiple pathways in placental development and function, genetic variations in the PPAR-γ gene might alter the risk of PE. A rare mutation in exon 6 of the PPAR-γ, Pro467Leu was reported in a woman who developed T2D, hypertension, and two pregnancies complicated by preeclampsia (Barroso et al., 1999). Although Pro12Ala was identified years ago, it has only been tested in relation to preeclampsia in a small genetic study in a Finnish population (133 cases and 115 controls), which showed no association between Pro12Ala and preeclampsia (Laasanen et al., 2002). The association between preeclampsia and this variant as well as other variants in PPAR-γ merit investigation in a larger study. This is not only in the view of the biological function of PPAR-γ but also

because PPAR-γ variants are reported as risk factors for T2D, CAD, and alteration in BMI, all of which are associated with preeclampsia.

In relation to birth weight Eriksson et al (2002) showed no association between birth weight and Pro12Ala in a relatively small study (470 participants) of Finnish subjects (Eriksson et al., 2002). The lack of association between birth weight and Pro12Ala was replicated in two larger studies (from Finland and Germany) that were adequately powered to detect down to 8% genotype attributable variation in birth weight (Bennett et al., 2008, Pfab et al., 2006). There was no association between Pro12Ala and SGA (171 SGA compared to 233 AGA) in a French population (Jaquet et al., 2002). Further investigation is required in this specific area, including testing of the maternal genotype at this locus.

Shaat et al showed no association between Pro12Ala and GDM in Scandinavians in a large study (Shaat et al., 2007). The same findings were reported in a Korean population (Cho et al., 2009). However Lauenborg et al (2009) included the Pro12Ala polymorphism in a study of 11 T2D variants in Danish women. The increased risk for GDM associated with the Ala allele was not significant (OR 1.16, 95% CI 0.82–1.52, P= 0.3), but when they tested 11 T2D variants, they found a highly statistically significant additive effect of multiple alleles on risk of the GDM (OR: 1.18· 95% CI 1.10–1.27, P= 3.2×10^{-6}), (Lauenborg et al., 2009).

In summary, cumulative evidence indicates the central role of a common alanine for proline substitution at codon 12 and C1431T in exon 6 of PPAR-γ in T2D, obesity, CAD, and atherosclerosis. This observation is of considerable interest

because patients with these disorders are at increased risk of pregnancy related disorders like GDM, and preeclampsia, and growth restriction in the offspring.

2.2.5 Cyclin dependent kinase inhibitor 2 B- (CDKN2B) antisense RNA 1 (CDKN2B-AS1)

2.2.5.1 Introduction

Cyclin dependent kinase inhibitor 2 B antisense RNA 1 (CDKN2B-AS1), also known as (Antisense Non-coding RNA in the INK4 Locus- (ANRIL) maps to chromosome 9p21, just downstream of the CDKN2A/2B genes. In 2008, Broadbent et al showed that CDKN2B-AS1 is expressed in many cell lines of vascular lineage including umbilical vein endothelial cells, microvascular endothelial cells, coronary smooth muscle cells, and in biopsies from carotid endarterectomy and abdominal aortic aneurysm, marking this gene as an important factor in vascular pathophysiology (Broadbent et al., 2008).

CDKN2B-AS1 has been classified as a long non-coding RNA (LncRNA). LncRNAs are empirically considered to be RNA molecules longer than 200 nucleotides that lack an open reading frame (ORF). They often overlap with, or are interspersed between, multiple coding and non-coding transcripts (Mercer et al., 2009). The LncRNAs have diverse origins. They may arise from frame shift disruption of a protein-coding gene, following a chromosomal rearrangement, duplication of a noncoding gene or neighbouring repeats. Lastly they can arise from insertion of a transposon element (Ponting et al., 2009). The LncRNAs regulate neighbouring protein-coding genes via cis-acting mechanisms, and target distant genes via trans-acting mechanisms. So, LncRNAs such as

CDKN2B-AS1 have been implicated in a variety of cellular pathways and may influence gene transcription through multiple mechanisms that might include RNA interference, gene silencing, chromatin remodelling, or DNA methylation (Mercer et al., 2009).

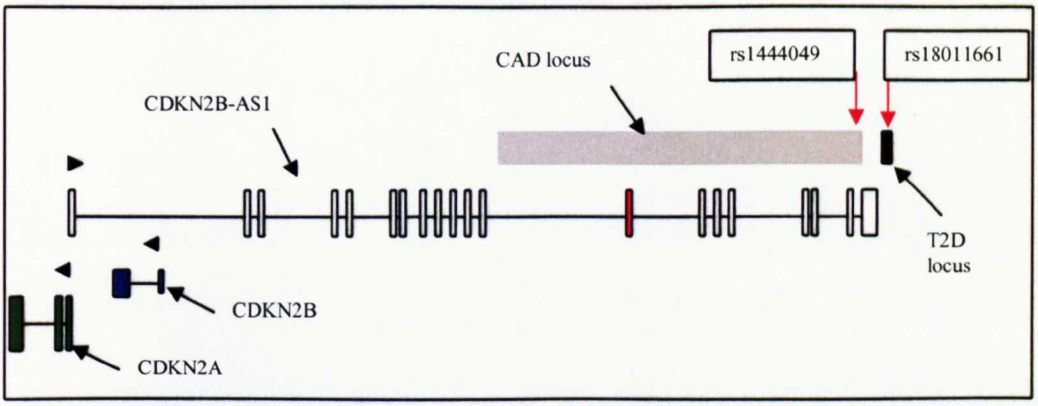
2.2.5.2 Genomic structure of CDKN2B-AS1

The CDKN2B-AS1 gene is formed of about 126.5 Kbp and maps to chromosome 9p21. It is formed of 20 exons of which exon 13 is alternatively spliced from the long intron 12. Four transcripts have been identified for this gene; the longest transcript is formed of exons 1-12 and 14-20. The transcripts, both long and short, have been detected in many tissues; of interest for this study, these include placenta. Exon 1 of the CDKN2B-AS1 gene is overlapping with exon 1 (5' end) of the cyclin dependent kinase inhibitor 2A (CDKN2A) gene. Also the entire CDKN2B gene is overlapping with the 1st intron of the CDKN2B-AS1 gene (Figure 2-5). CDKN2A/and CDKN2B genes are implicated in the regulation of cell cycle and senescence.

Exons of this gene vary in their length between 670 bp down to 70 bp and are formed of multiple simple repeats. Conserved non coding regions have been identified within primates, and less conserved sequences were noticed with orthologous species, indicating that this gene has evolved recently in mammals with new functions or underwent marked evolutionary changes with degenerate function (Jarinova et al., 2009).

Figure 2-5: Schematic presentation of chromosome 9p21 region

The CDKN2B-AS1 gene showing exons (open rectangles) with the alternatively spliced exon 13 (red), and CDKN2A/B genes in green and blue respectively. CAD and T2D loci are shown in grey in black rectangles respectively. The black arrows indicate the direction of transcription. The red arrows refer to SNPs included in this study, rs1333049 (right) and rs10811661 (left).



2.2.5.3 Polymorphisms of CDKN2B-AS1

Genetic polymorphism in this region has been consistently associated with coronary artery disease (CAD) in European populations, with suggested a population attributable risk of around 20% for heterozygous and 30% for homozygous carriers of the risk allele (McPherson et al., 2007).

In a population of European descent from Canadian and US residents McPherson et al (2007) carried out a genome wide study testing 100,000 SNPs for association with CAD, which included a group of SNPs located on chromosome 9p21. From within this group the two SNPs which gave the strongest signals for association with CAD were rs10757274 (OR 1.29, 95% CI 1.09–1.52, P=0.01) and rs2383206 (OR1.26, 95% CI 1.09–1.46, P=0.04). These two SNPs are 20 kb apart and in strong linkage disequilibrium (r^2 = 0.89) (McPherson et al., 2007).

The CAD risk locus on chromosome 9 is represented by a haplotype spanning approximately 58 kb carrying 37 highly correlated SNPs from rs10757269 to rs1333049. The region is flanked by 2 recombination hotspots with no evidence of long range linkage disequilibrium (LD) on chromosome 9, consequently the causative allele most likely resides within this region (McPherson et al., 2007, Broadbent et al., 2008).

Further studies including thousands of CAD cases and controls from British, Dutch, Italian, and US-Icelandic populations replicated the findings of McPherson et al, with the C allele of rs1333049 SNP (MAF of the C allele 0.45) showing the strongest and most reproducible association (Gori et al., 2010, Samani et al., 2007, Helgadottir et al., 2007). This locus was also reported as a risk factor for CAD in the WTCCC GWAS (OR 1.47, 95% CI 1.27-1.7, $P=1.16x10^{-13)}$ (WTCCC., 2007). A meta-analysis including 12004 cases and 28949 controls from 7 European case control studies replicated the strong association of the C allele of rs1444049 SNP and CAD, (OR 1.24, 95% CI 1.20-1.29, $P= 6.04x10^{-10}$) (Schunkert et al., 2008).

The data on gene environment interaction at this locus is not conclusive. The risk conferred by this genetic variant for CAD was not altered when adjusting for age, gender, smoking, plasma lipid levels, blood pressure, diabetes, and plasma C-reactive protein concentrations, indicating these factors did not modify the genetic effect (Broadbent et al., 2008). In contrast, Doria et al (2008) reported that the magnitude of effect of the CAD locus on chromosome 9 was doubled in T2D patients and the risk was doubled again in patients with poor glycaemic

control (defined by average Haemoglobin A1 in the 9 years before study entry) (Doria et al., 2008).

It has been noted that the association between these SNPs and CAD was stronger in patients with CAD who had not experienced myocardial infarction (MI), excluding this gene from being a thrombosis or plaque rupture gene (Helgadottir et al., 2007). The risk allele was associated with abdominal aortic and intracranial aneurysm and stroke independently of CAD, and it has therefore been proposed that the risk allele increases the atherosclerotic process (Helgadottir et al., 2008). It has been suggested that changes in the collagen and/or elastin content of the vessel wall may be the underlying mechanism for the association of this region with arterial diseases (Bjorck et al., 2009).

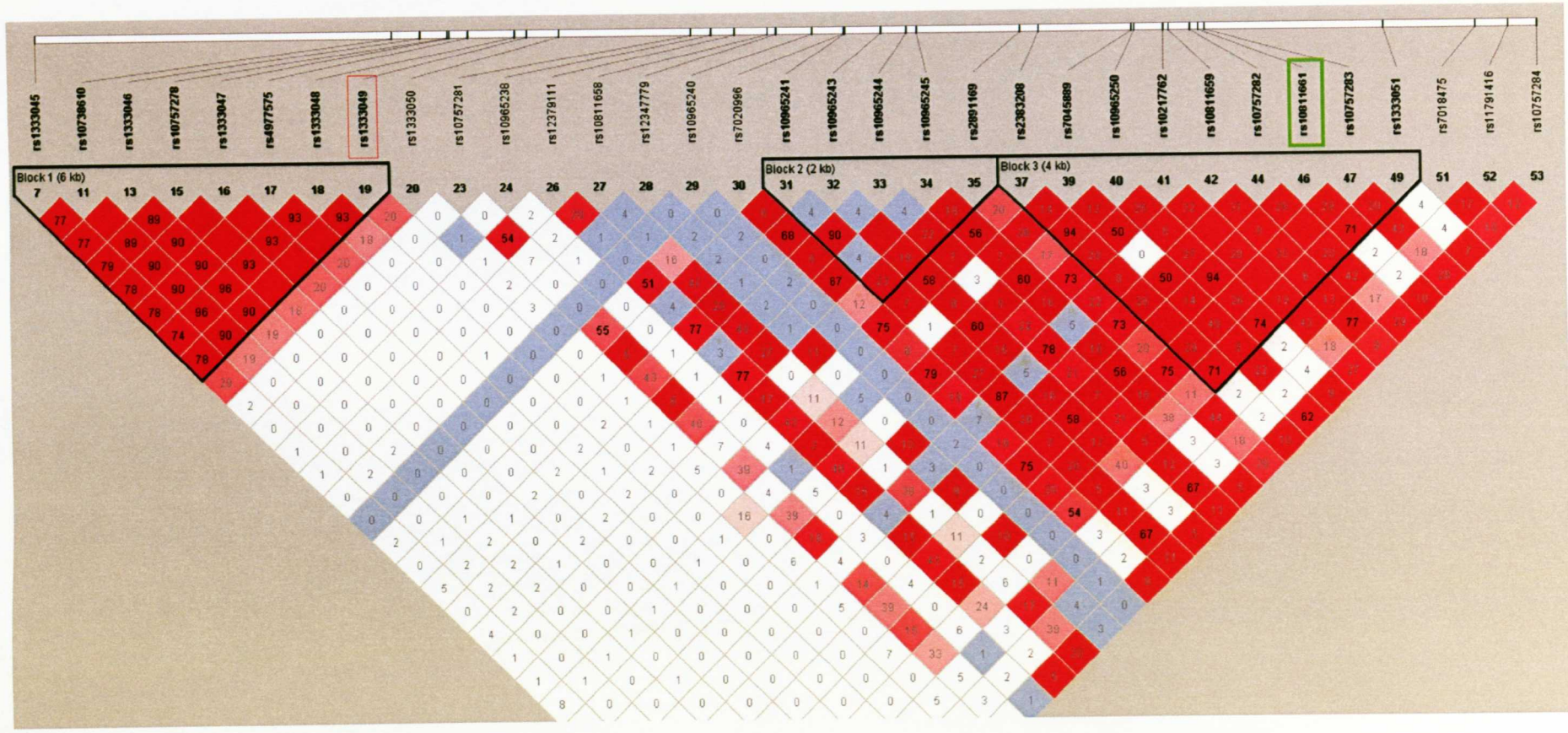
The same region of chromosome 9p21 was repeatedly associated with T2D in GWAS in a UK population (WTCCC., 2007), Scandinavians (Saxena et al., 2007), and in Finns (Scott et al., 2007). The T allele (major allele frequency of 80% in the European population) of rs10811661 has been shown as the strongest and most consistently associated SNP in a meta-analysis of T2D GWAS, (OR 1.19, 95% CI 1.11-1.28, $P=4.9x10^{-7)}$ (Zeggini et al., 2007). The association of this marker with T2D was also replicated in a case control study in French subjects (OR 1.43, (95% CI 1.24-1.64 0 $P= 3.8 \times 10^{-7}$) (Duesing et al., 2008).

Many pathological features of T2D and the metabolic syndrome are predisposing factors for atherosclerosis and CAD. For example, increased oxidative stress and hyperlipidaemia associated with T2D and metabolic syndrome lead to endothelial damage and atheroma formation. It is therefore potentially the same

chromosomal locus which has been associated with both CAD and T2D, suggesting joint genetic and pathophysiological mechanisms. Nevertheless T2D SNPs at chromosome 9p21 are different from those of CAD. The rs10811661 SNP is located 10 kbp downstream of the haplotype reported as a CAD susceptibility locus (Figure 2-6). Analysis of Hap Map data showed that rs10811661 and rs1333049 are located in different LD blocks (Figure 2-6). A simultaneous test of CAD and diabetes susceptibility with CAD and T2D-associated SNPs indicated that these associations were independent of each other (Broadbent et al., 2008, Gori et al., 2010). Furthermore, none of the quantitative traits of the metabolic syndrome have been associated with the CAD locus.

Figure 2-6 : The pattern of LD in chromosome 9p21 region

CAD and T2D LD regions in Europeans are clearly separate. The SNPs in the CAD region are highly correlated. The two SNPs that have been identified as most strongly associated with CAD (rs1333049, red box) and T2D (and rs10811661, green box) are shown.



2.2.5.4 Functional effects of the CAD risk variant of CDKN2B-AS1

The association between the 9p21.3 locus and CAD and T2D risk is robust and investigation of the molecular basis for a causal relationship is currently ongoing. The high-risk CAD haplotype overlaps exons 12 to 20 of CDKN2B-AS1 that is expressed in many cell types known to be affected by atherosclerosis (Gonzalez-Navarro et al., 2010).

Jarinova et al (2009) examined four conserved non-coding regions of chromosome 9p21 for enhancer activity using a luciferase reporter gene. A region termed CNS3 (chr9:22108881–22111116) showed an enhancer activity compared to the other 3 regions within the 58kbp CAD locus. The CNS3 sequence cloned from a homozygous carrier of the C of rs1333045 (risk allele) exhibited 73% higher enhancer activity compared to the same sequence cloned from a reference homozygote in aortic smooth muscle cells. SNP rs1333045 was in high linkage disequilibrium with rs1333049 (r^2 0.86) that repeatedly associated with CAD (Jarinova et al., 2009). This enhancer activity might affect the expression of the CDKN2B-AS1 gene itself or it might affect the expression of other neighbouring genes, CDKN2A/B as well as distant genes (upstream and downstream). Interestingly the SNP rs1333049 was the most significantly associated SNP in relation to CAD in a 100 Kb susceptibility region identified in the WTCCC (2007) study.

Deletion of the orthologous 70-kb non-coding interval on mouse chromosome 4 results in enhanced proliferation properties of vascular cells, increased tumour incidence, and increased weight of the knock out mice (Visel et al., 2010).

Being implicated in T2D and CAD, this locus on chromosome 9p21 would be of interest for investigation in preeclampsia and FGR.

2.2.6 Potassium Inwardly-Rectifying Channel, Subfamily J, Member 11, KCNJ11

2.2.6.1 Introduction

Fifteen inwardly rectifying potassium (Kir) subunit genes have been identified and classified into seven subfamilies (Kir1.x to Kir7.x). These subfamilies are categorized into functional groups: 1) classical Kir channels (Kir2.x); 2) G-protein-gated Kir channels (Kir3.x); 3) ATP-sensitive channels (Kir6.x), to which KCNJ 11 belongs; and 4) K-transport channels (Kir1.x, Kir4.x, Kir5.x, and Kir7.x). They share the basic Kir subunit, discussed later (reviewed in (Hibino et al., 2010).

ATP-sensitive channels are proposed to link the cellular metabolism and electrophysiology of the cell membrane. They are expressed in different tissues throughout the body, including pancreatic alpha and beta cells, cardiac myocytes, vascular smooth muscles, skeletal muscles and in neurons. These channels are responsible for regulating diverse processes including cellular excitability, vascular tone, heart rate, and insulin release.

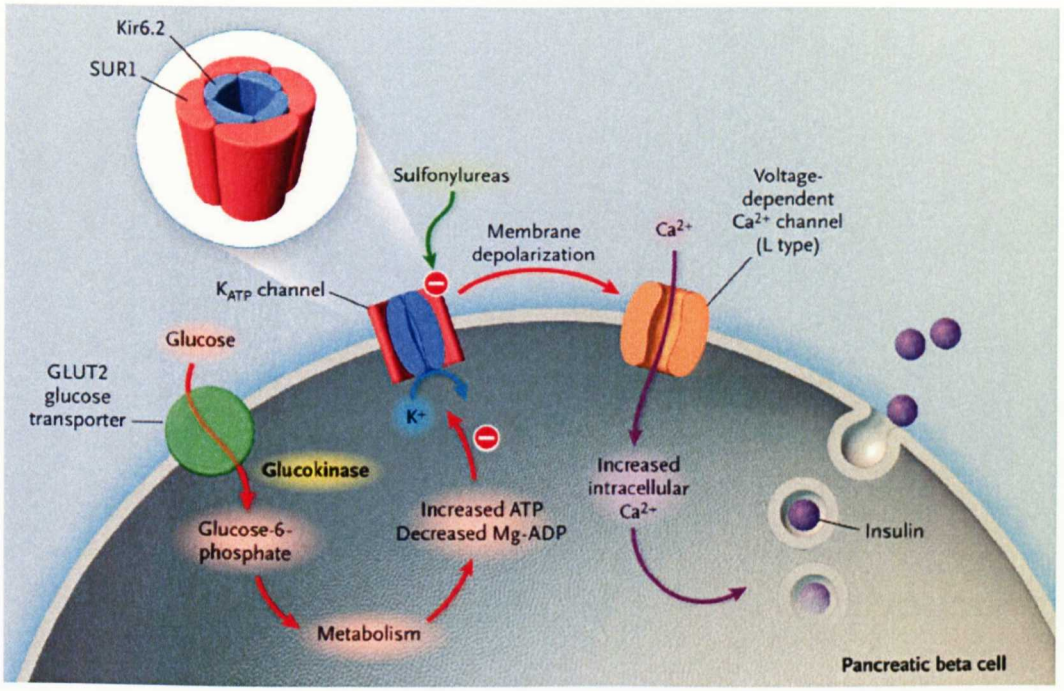
The KATP channel is formed of two different subunit types arranged as a hetero-octamer: 4 Kir6.x subunits located at the centre (pore forming) and 4 sulfonylurea receptor subunits (SURx) located at the periphery, (Figure 2-7); (Shyng and Nichols, 1997). Subunits cannot reach the cell membrane alone as

they carry an endoplasmic reticulum retention signal, which needs to be masked by the other subunits to guarantee that only fully functional KATP channels are trafficked to the surface membrane. Different combinations of Kir6.x and SURx in different tissues have been identified, based on both mRNA and protein expression data. The KATP channel is formed of SUR1+Kir6.2, SUR2A+Kir6.2, SUR2B+Kir6.1, and SUR2B+Kir6.2 in pancreatic b-cells, cardiac muscle, vascular smooth muscle, and non-vascular smooth muscle channels, respectively (Olson and Terzic, 2010).

ATP is the main regulator of the KATP channel; it closes the channel and maintains its activity in the presence of Mg^{+2}. In pancreatic b-cells for example, when serum glucose levels are high, abundant cellular ATP leads to closure of the KATP channel. As a result K^+ influx into the cell is diminished, and the cell membrane undergoes depolarization. This is followed by opening of the Ca^{+2} channels with increase in cytoplasmic Ca^{+2} and subsequent exocytosis of insulin (Figure 2-7). KATP channels in b-cells regulate insulin secretion by linking the metabolic state of the cell to membrane potential (reviewed in (Hibino et al., 2010).

Figure 2-7: Structure of KATP channel in the b-cell of the pancreas

The role of KATP channel in insulin secretion as well as the effect of sulfonylurea drugs, adapted from (Gloyn et al., 2004), with permission.

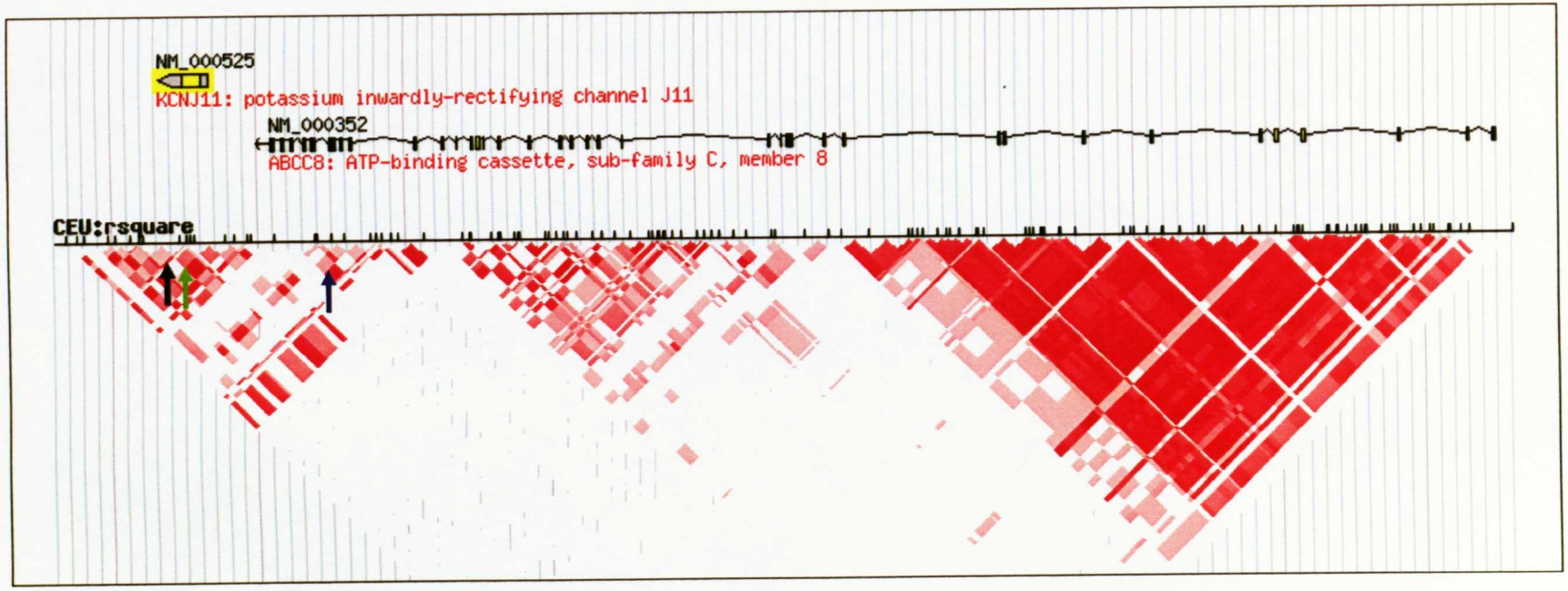


2.2.6.2 Genomic structure and polymorphism of the KCNJ11

The KCNJ11 (Kir6.2) gene is mapped to chromosome 11p. It is formed of one exon encoding 390 AA. This gene is located approximately 4.5 Kb away from the 3' end of the ABCC8 (SUR1) gene. The ABCC8 gene encodes the second structural part of the KATP channel, spans more than 84 kb of genomic DNA and encodes a 1,582 amino acid protein from 39 exons. The entire KCNJ11 gene and the 3' end of the ABCC8 gene are enclosed in single LD block. This LD block harbours a group of highly correlated SNPs, some of which are coding SNPs and designated as T2D risk variants (Figure 2-8).

Figure 2-8: LD blocks and SNPs in both KCNJ11 (Kir6.2) and ABBC8 (SUR1) genes

The entire KCNJ11 and the 3' part of the ABCC8 gene are contained in LD block of 20Kb carrying 38 genotyped SNPs, 5 of which are coding SNPs. Rs5215 (I337V, black arrow) and rs5219 (E23K, green arrow) in KCNJ11, and Rs 757110 (S1369, blue arrow) in ABCC8 are shown (created by Hap Map # 24).



2.2.6.3 Role of KCNJ11 variants in type 2 diabetes

KCNJ11 was linked to the pathology of T2D based on three factors: the association of rare inactivating mutations of the gene with hyperinsulinaemic hypoglycaemia of infancy; the association of activating mutations with neonatal diabetes (Gloyn et al., 2004, Meissner et al., 1999); and the successful use of sulfonylureas which target the KATP channels in the treatment in T2D (Mannhold, 2004).

Conflicting early reports about the association of KCNJ11 variants with T2D were probably due to the small sample sizes. However, in a large UK study including 850 T2D cases and 1200 controls, Gloyn et al (2003) demonstrated that the K allele of E23K (rs5219) is a risk variant for T2D (OR 1.18, 95% CI 1.04 –1.34, P= 0.01), (Gloyn et al., 2003). They also carried out a meta-analysis including 5 earlier studies, and confirmed the association of this variant with T2D (OR 1.23, 95% CI 1.12–1.36, P= 0.000015).

In a study of more than 3000 Scandinavians, Florez et al (2004) replicated the association of E23K with type 2 diabetes (OR 1.17, 95% CI 1.05–1.32, P= 0.003). The authors indicated that E23K along with S1369A (rs757110), (r^2> 0.9) in a large LD block representing KCNJ11 and the 3' end of the ABCC8 gene, are the true signals for the association. Testing multiple haplotypes, using 15 tagging SNP in this LD block, gave no signals stronger than the E23K variant for the association with T2D (Florez et al., 2004).

Replication in different populations including Korean (Koo et al., 2007), Japanese (Doi et al., 2007) and Russian (Chistiakov et al., 2009) confirmed a true

association. Two markers, rs5215 (I337V) and rs5219 (E23K), were associated with T2D in GWAS for T2D (WTCCC., 2007).

Prospective studies were also carried out to test the role of this variant on the conversion from normal glucose tolerance (NGT) to impaired glucose tolerance (IGT) and T2D and also to assess the interaction of this variant with lifestyle intervention and medication on the conversion to glucose intolerance. Laukkanen et al (2004) was the first to test such a hypothesis using around 450 IGT Finnish subjects with 3 years follow up. This study indicated that the E23K alone was not associated with conversion from IGT to T2D (Laukkanen et al., 2004). When combined with other variants from within the ABCC8 gene it increased the risk of conversion to 6-fold in comparison to the 2-fold increase in risk conferred by the ABCC8 genetic variants alone. A more recent study by Florez et al (2007) which included 3500 subjects with IGT followed up for 3 years, reported that the K allele of E23K was associated with a lower risk of developing T2D in the placebo and in the life style intervention groups (Florez et al., 2007). These findings suggested that the K allele might be associated with the earlier stage of b-cell dysfunction (from normal glucose tolerance (NGT) to IGT) rather than the deterioration of b-cell function leading to conversion from IGT to T2D. It is possible that the use of IGT and obese subjects who carry a baseline genetic and environmental risk of T2D might weaken or obscure the actual association of conversion from IGT to T2D. This was not the case in the study carried out by Lyssenko et al (2008) that included more than 15000 NGT subjects with 23 years follow up in which the K allele was associated with increased risk of conversion from both NGT and IGT to T2D (Lyssenko et al., 2008).

Both *in vivo* and *in vitro* studies showed that the E23K SNP is functional, with the K allele causing lower insulin secretion. Neilson et al (2003) reported that in a group of 519 NGT individuals, the K allele was associated with lower insulin secretion in response to OGTT testing (P =0.022, adjusted for BMI), (Nielsen et al., 2003). Florez et al (2004) reported that a 20% -30% reduction in insulin release was explained by genotype at E23K when tested in 500 NGT individuals and 39 sibling pairs (Florez et al., 2004).

In agreement with *in vivo* studies, Schwanstecher et al (2002) were the first to examine the mechanism by which E23K variants lead to reduced insulin secretion. They suggested that the K allele, located at the transmembrane domain of the subunit and not directly related to ATP binding of the receptor, was associated with reduced ATP sensitivity of the Kir6.2/SUR1 channel complex. This results in increased probability of an open KATP channel, and subsequent suppression of insulin secretion (Schwanstecher et al., 2002). Moreover failure to inhibit glucagon release by alpha cells after glucose stimulation was suggested as an alternative mechanism by which this polymorphism leads to glucose intolerance (Tschritter et al., 2002).

Although the E23K variant of KCNJ11 is the strongest T2D candidate within the LD block, which includes KCNJ11 and the 3' end of ABCC8 gene, the impact of the neighbouring ABCC8 gene cannot be ruled out, in particular the non-synonymous S1369A substitution. This polymorphic site is also reported to reduce the KATP channel activity (Hamming et al., 2009).

Because the KATP channel of the pancreatic b-cells is a target for sulfonylurea drugs as a treatment for T2D, testing the effects of genetic variants encoding this receptor on response to treatment was of interest. Sulfonylurea drugs interact with SUR1 and KIR6.2 to inhibit K influx through the receptor leading to membrane depolarization with subsequent opening of Ca^{+2} channels that stimulate insulin secretion. An initial report by Glyon et al (2001) using a cohort of 364 newly diagnosed diabetic UK patients did not find any significant association of the E23K variant with the response to sulfonylurea therapy (Gloyn et al., 2001). A second study by Sesti et al (2006) using 552 patients, indicated that the K allele was more common in those who failed to respond to sulfonylurea or a combination of sulfonylurea and metformin treatment than those who responded to this treatment (OR 1.45, 95% CI 1.01–2.09; P= 0.04) (Sesti et al., 2006).

2.2.6.4 Role of KCNJ11 variants in pregnancy disorders

Because insulin is a growth regulator, Weedon et al (2003) tested the effect of the E23K variant in relation to birth weight. They showed that the fetal genotype at E23K was not associated with reduced birth weight in about 1000 UK subjects and suggested that it does not affect insulin secretion in utero (Weedon et al., 2003). In support of this Bennett et al (2008) showed the same lack of effect in a large Finnish cohort of 5000 individuals (Bennett et al., 2008). On the other hand this variant was reportedly associated with GDM (OR 1.17, 95% CI 1.02–1.35; p=0.027) in a large study of Scandinavian women (Shaat et al., 2005)and was marginally associated in a small study of Danish women (OR 1.20, 95% CI 0.99 –1.45, P= 0.07), (Lauenborg et al., 2009). Both of these studies are in line with the notion that b-cell dysfunction is a key factor in the development of GDM.

KCNJ11 is an important T2D candidate gene, and merits investigation in relation to preeclampsia and FGR.

2.3 Aim of the study

1- To test the shared genetic basis between T2D and preeclampsia in white Caucasian women from the UK by comparing genotype and allelic frequencies of 11 SNPs in 5 T2D susceptibility genes in normotensive pregnant women and in women with preeclampsia.

2- To investigate the effect of maternal genotype at these 11 loci in 5 T2D susceptibility genes on the risk of having growth restricted pregnancies.

3- To determine whether genotype at these loci in women with healthy pregnancy are associated with their own BMI and/or birth weight of their babies, in Caucasian women from the UK.

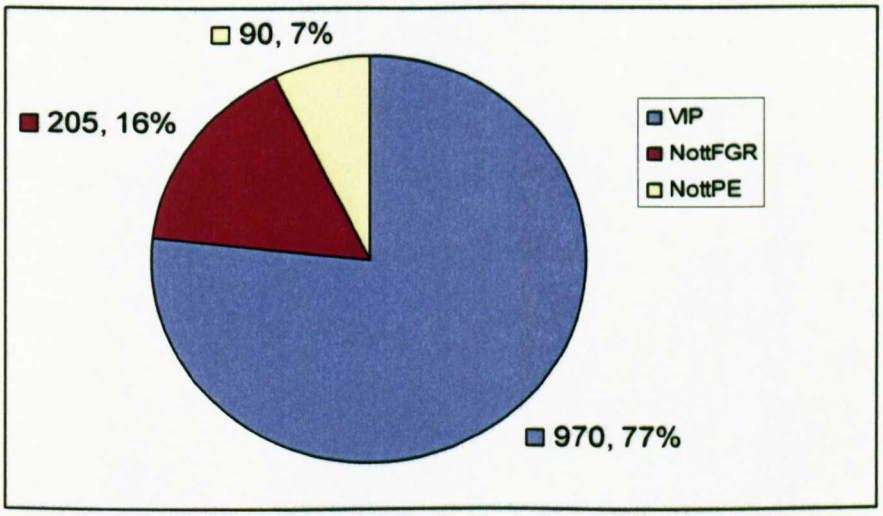
2.4 Methods

2.4.1 Patient recruitment

Our cases and controls were recruited from three different collections which are described below. The numbers from each study are shown in (Figure 2-9).

Figure 2-9: Number and frequency of participant women from different collections

VIP: Vitamins in preeclampsia, NottFGR: Nottingham FGR collection, NottPE: Nottingham preeclampsia collection.



2.4.1.1 Vitamins in Preeclampsia (VIP) trial

In this study the role of the antioxidant vitamins C and E in the prevention of preeclampsia were investigated in a randomised controlled trial to which women with known risk factors for preeclampsia from 25 UK hospitals in ten geographical areas were recruited from 2003 to 2005 (Poston et al., 2006).

Women were eligible for recruitment for the VIP study if they were pregnant at gestational age 12-20 weeks with one or more factors known to increase risk for preeclampsia. They included preeclampsia in the pregnancy preceding the index pregnancy requiring delivery before 37 weeks of gestation; a diagnosis of HELLP syndrome and/or eclampsia in any previous pregnancy at any stage of gestation; diastolic blood pressure of 90 mm Hg or more at or before 20 weeks gestation in the current pregnancy; multiple pregnancy; and abnormal uterine artery Doppler waveform. Women with the following systemic risk factors were also eligible: hypertension requiring medication currently or previously; pre-gestational diabetes requiring insulin or oral hypoglycaemic therapy; anti-phospholipid syndrome; chronic renal disease; primiparity with BMI at the first antenatal appointment of 30 kg/m^2 or more. Ethical approval from The South East Multi Regional Research Ethics Committee and written informed consent from the participants were obtained.

The results of the original study showed that preeclampsia and SGA, the primary and secondary outcomes in this study, were unaffected by the use of 1000 mg vitamin C and 400 IU vitamin E daily from the second trimester of pregnancy

until delivery. 970 women representing 49% of the white western European women from the VIP study were included in this genetic study (Figure 2-9).

From the VIP collection 343 and 574 women were included as preeclampsia cases and controls respectively, of which 49 and 63 % were presented with one of the above mentioned risk factors. 110 and 471 women were considered as FGR cases and controls respectively, of which 56 and 53 % were presented with one of the aforementioned risk factors. But when BMI > 30 (kg/m^2) in primiparous women was considered as a risk factor 100% of the VIP particepants were considered as women with risk factors. This methos of recruitment was the basic and major difference between participants from the VIP study and the other 2 studies. Clinicopatholigical features are presented in Table 2-2.

2.4.1.2 Nottingham FGR collection

This collection comprised DNA samples from volunteers recruited between 2000 and 2002 from Nottingham City Hospital and the Queen's Medical Centre, Nottingham after approval of the ethics committees of both hospitals. Written consent was obtained from participants who were self identified as being of white western European descent.

FGR cases were identified antenatally when ultrasound scan estimated the abdominal circumference $\leq 5^{th}$ centile for gestational age using locally derived charts, and confirmed postnatally by individualized birth weight centile $\leq 5^{th}$ centile using GROW software (http://www.gestation.net/register/growreg/grow_reg.asp). Controls were chosen from women undergoing Caesarean section, induction of labour for causes not related to fetal growth, or normal

delivery, with babies' birth weight above the 10th centile and below the 96th. 205 participants from this collection were included in the present study (Figure 2-9).

All women were normotensive at booking, at less than 20 weeks gestation dated by ultrasound scan, had no medical problems (other than pregnancy induced hypertension), and had a structurally and chromosomally normal singleton infant. Pregnancies complicated by infection, anaemia and drug use were excluded. From the NottFGR collection 99 mothers were included as controls for both the FGR and preeclampsia studies, 106 mothers were included as FGR cases, of whom 15 also developed preeclampsia, as defined by the ISSHP. Demographic and clinical data are presented in Table 2-2.

2.4.1.3 Nottingham preeclampsia collection

Ninety women (78 nulliparous) from the Nottingham preeclampsia collection were included in this study (Figure 2-9). They comprised previously healthy women with preeclampsia, recruited during pregnancy for genetic case control studies between 1993 and 1996. Preeclampsia was defined as a diastolic blood pressure exceeding 90 mmHg and a systolic blood pressure exceeding 140 mmHg in a previously normotensive woman, which returned to normal within three months of delivery, and was associated with proteinuria in excess of 300mg /L in a 24 h collection or two pluses on testing a random sample with Multistix (Bayer, Newbury, UK). Clinical and demographic data are presented in Table 2-2

Table 2-2: Clinical and demographic data for participants in preeclampsia and FGR genetic case control studies stratified by collection

VIP: Vitamins in pregnancy, NottFGR: Nottingham FGR collection, NottPE: Nottingham preeclampsia collection. * Data are presented as median (interquartile range) for continous data and number and percentage for categorical data. In preeclampsia study, data on smoking habits were available for 99 and 98 % of cases and controls of the VIP women and 93 % of NottPEcollection. In FGR study, data on smoking habits were available for 98 and 97 % of cases and controls of the VIP women and 93 % of the NottPE. ** 100% of the VIP participants carried at least one risk factor if BMI > 30 (kg/m^2) in primiparous women was considered as a risk factor.

I- Preeclampsia study

	VIP		NottFGR		NottPE
	Cases N=343	Controls N=574	Cases N=15	Controls N=99	Cases N=90
Maternal age	32 (28-36)	31 (26-35)	25 (19-29)	29 (20-34)	28 (24-32)
Nulliparous	81(24)	402 (70)	13 (86)	56 (56)	78 (87)
BMI(kg/m^2)	27 (24-33)	31 (24-35)	22 (20-26)	24 (22-27)	25 (23-28)
Smoking habits*					
Never smoker	189 (55)	291 (52)	6 (40)	74 (74)	70 (83)
Previous smoker	111 (33)	183(33)	3 (20)	11 (11)	2 (2.5)
Current smoker	39 (12)	86 (15)	6 (40)	14 (14)	12 (14)
Systolic blood pressure (mmHg)	142 (120-160)	120 (110-130)	160 (150-172)	110 (101-120)	160 (150-166)
Diastolic blood pressure (mmHg)	95 (91-106)	70 (64-80)	101 (98-104)	65 (60-70)	100 (97-106)
Gestational age at delivery (days)	268 (258-275)	273 (263-285)	224 (189-238)	273 (266-280)	259 (238-273)
Fetal Weight (gm)	3140 (2550-3580)	3345 (2940-3760)	1095 (670-1530)	3370 (3120-3660)	2545 (1885-3140)
Risk factors**	167 (49)	359 (63)	-	-	-

II-FGR study

	VIP		NottFGR		NottPE
	Cases N=110	Controls N=471	Cases N=106	Controls N=99	Cases N=27
Maternal age	32 (28-36)	31 (26-35)	27 (21-30)	29 (20-34)	27 (24-32)
Nulliparous	50 (46)	343 (73)	65 (61)	56 (56)	18 (67)
BMI(kg/m2)	29 (24-35)	32 (27-36)	23(21-25)	25 (22-27)	25 (21-28)
Smoking habits*					
Never smoker	64 (59)	237 (52)	47 (45)	74 (74)	23 (92)
Previous smoker	29 (27)	150 (33)	10 (9)	11 (11)	1 (4)
Current smoker	15 (14)	72 (15)	49 (46)	14 (14)	1 (4)
Systolic blood pressure (mmHg)	149 (130-170)	120 (110-130)	110 (104-124)	110 (101-120)	160 (155-171)
Diastolic blood pressure (mmHg)	99 (77-110)	72 (68-80)	65 (60-72)	60 (60-70)	105 (99-110)
Gestational age at delivery (days)	262 (240-275)	277 (268-286)	259 (224-266)	273 (266-280)	238 (224-259)
Fetal Weight (gm)	2210 (1690-2565)	3356 (3180-3860)	2110 (1377-2472)	3370 (3120-3660)	1720 (1120-2180)
Risk factors **	62 (56)	251 (53)	-	-	-

2.4.2 Preeclampsia and fetal growth restriction phenotype definition

From within the three collections a unifying definition of preeclampsia that fulfils the ISSHP criteria was applied. Preeclampsia was defined as diastolic blood pressure of ≥90 mm Hg and/ or systolic blood pressure of ≥140 mm Hg measured twice at least 4 h and up to 7 days apart at or after the 20th week of gestation, associated with proteinuria of ≥300 mg protein over 24 hours or two readings of 2+ or more on dipstick analysis of mid-stream urine. In women in the VIP study with hypertension before pregnancy or before 20th weeks of gestation the development of new proteinuria was required for the diagnosis of preeclampsia. In women with pre-existing proteinuria diastolic blood pressure of 90 mm Hg or more occurring after 20 weeks of pregnancy was required for the diagnosis of preeclampsia. HELLP syndrome and eclampsia were included as cases only if they were hypertensive and proteinuric.

Fetal growth restriction (FGR) is defined as a singleton baby with birth weight ≤5th centile using an individualized birth weight centile (correcting for maternal height, weight, and parity, and fetal gestation and sex). Ninety nine (22.1%) of preeclampsia cases were also FGR.

Controls for preeclampsia were lacking all of the above criteria in the current pregnancy and did not suffer from preeclampsia, eclampsia, or HELLP in any previous pregnancies. Controls for the FGR study were defined as a singleton baby with birth weight ≥10th and ≤90th centile using individualized birth weight centile, excluding mothers with preeclampsia, eclampsia, or HELLP in the current or previous pregnancy (s). data are presented in Table 2-2

represents how many cases and controls were included from the 3 collections.

2.4.3 Power calculation

The preeclampsia study included 448 cases and 673 controls. The FGR study included 243 mothers with pregnancies complicated by FGR and 570 controls. For an average MAF of 0.3, this study had a power of 80% to detect a genotype relative risk of 1.3 for preeclampsia and 1.4 for FGR at a P value of 0.05 assuming an additive model of inheritance. Detailed power calculations are presented in (Table 2-3). Quanto software V1.2.3 was used for this calculation (http://hydra.usc. edu/ gxe/) (Gauderman, 2002).

Table 2-3: Power calculations for case control studies

These calculations were based on a sample size of 673 controls and 448 cases for preeclampsia and 570 controls and 243 cases for the FGR study. The calculation assumed an additive model of inheritance, at a P value of 0.05. Percentage power is shown outside brackets for preeclampsia, and inside brackets for FGR.

Allele frequency	Genotype relative risk					
	1.1	1.2	1.3	1.4	1.5	2
0.1	10 (8)	25 (18)	47 (33)	69 (50)	85 (67)	99.9 (99)
0.2	14 (11)	40 (28)	70 (52)	89 (77)	97(88)	99.9 (99)
0.3	17 (12)	49 (34)	80 (61)	95 (82)	99 (93)	99.9 (99)

2.4.4 DNA extraction

DNA extraction from VIP samples was carried out using the Wizard® Genomic DNA Purification Kit from Promega according to the manufacturer's protocol (www.promega.com). Wizard® Genomic DNA Purification is based on a four-step process. The process starts with a cell lysis which ruptures red and white blood cells using Cell Lysis Solution. The second step is lysis of the nuclei of the white blood cells using the Nuclei Lysis Solution. This is followed by salt

precipitation of the proteins and cell debris in the Protein Precipitation Solution. The supernatant is then treated with isopropanol to precipitate DNA, followed by a further washing step using 70% ethanol. DNA is then suspended in the rehydration solution provided with the kit. The concentration was measured by absorbance at 280 nm using a Nanodrop instrument (ND-1000 V3.7.1; Nanodrop 1000 spectrophotometer; Thermo Scientific). The average yield from a 4-5 ml blood sample was 150 μg of high quality DNA which was stored at -70 °C.

Archived DNA from the FGR collection was previously extracted from leukocytes using the QIAamp blood maxi kit (Qiagen, Crawley, UK; www.qiagen.com). Stored DNA from the Nottingham preeclampsia collection was previously extracted from peripheral leucocytes using Nucleon II extraction kits (Scotlab, UK).

2.4.5 Genotyping

Genotyping was carried out by Kbiosciences using their in house developed technology, Kbiosciences Competitive Allele Specific PCR (KASPr), described in detail below. DNA samples were plated as 50μl of 20ng/μl for each well in deep v-shaped bottom 96 well plates, sealed and sent on dry ice for genotyping. Samples sent for genotyping by Kbiosciences included 970 samples from the VIP study, 205 maternal samples from the Nottingham FGR collection, 90 samples from the Nottingham preeclampsia collection of DNA. 78 samples were chosen at random and included as duplicates. Two no template controls (NTC) per 96 well plate, comprising PCR grade water were also included. For quality

control purposes, 2 SNPs were also genotyped in a subset of the samples in Nottingham laboratory using mutagenically separated PCR (MS-PCR).

2.4.5.1 The Kbioscience Competitive Allele Specific PCR (KASPr)

The KBioscience Competitive Allele Specific PCR genotyping system (KASPar) is a fluorescent-based genotyping technology (Figure 2-10).

The KASPr genotyping system comprises two components:

(1) The SNP specific assay consists of two allele specific forward primers, each with a unique unlabelled tail sequence at the 5' end, and one common (reverse) primer.

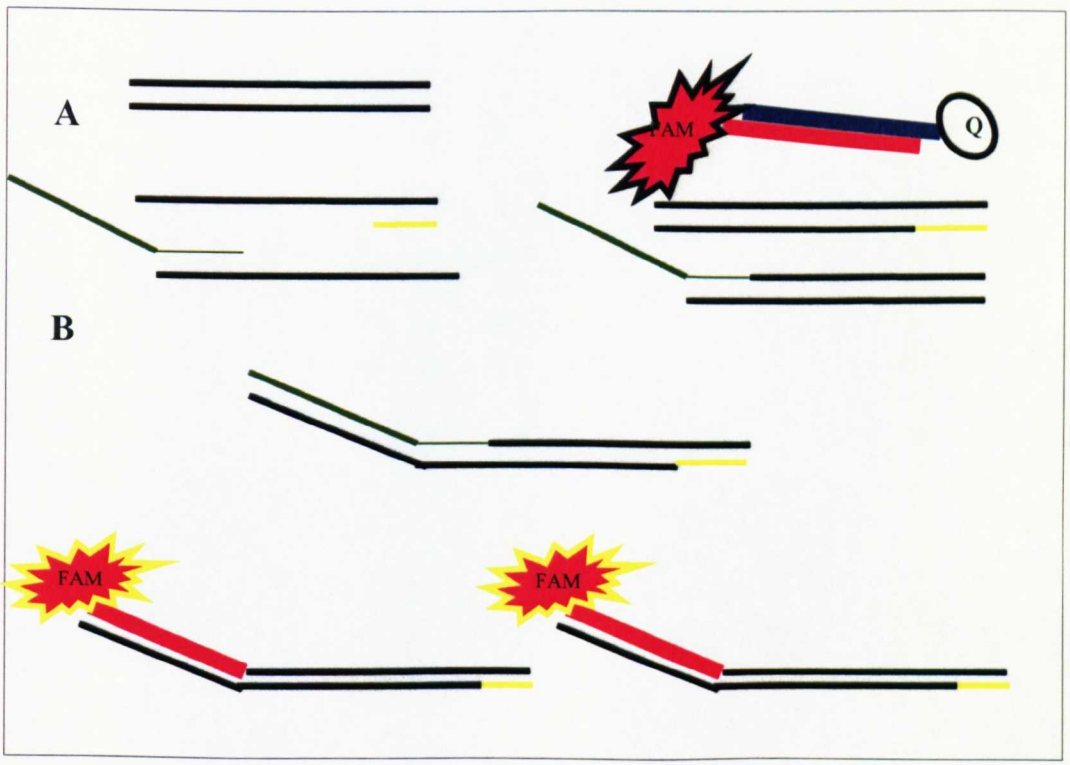
(2) The universal Reaction Mix, which contains specially developed Taq polymerase, two 5' fluorescently labelled oligonucleotides; one labelled with FAM and one with CAL Fluor Orange 560 one specific for each allele, two oligonucleotides with quenchers bound at the 3' ends. These oligonucleotides sequences are complementary to each other and to sequences of the tails of the allele specific primers. ROX is also included in the reaction mixture to be used as a reference dye that allows data to be normalised for the variation in the liquid volume and plate reader.

In the early stages of PCR, the quenched oligonucleotides bind their fluorescently labelled complementary oligonucleotides and all fluorescent signals are quenched. In the later stages of PCR, the fluorescently labelled oligonucleotide, corresponding to the amplified allele also gets incorporated into the template, and hence the fluorescence is no longer quenched. The generated

fluorescent signal can be detected by a FRET (Fluoresence resonance energy transfer) reader (http://www.kbioscience.co.uk/reagents/KASP.html). If the genotype at a given SNP is homozygous, only one or other of the possible fluorescent signals will be generated. If the individual is heterozygous, the result will be a mixed fluorescent signal.

Figure 2-10: KASPar genotyping technology

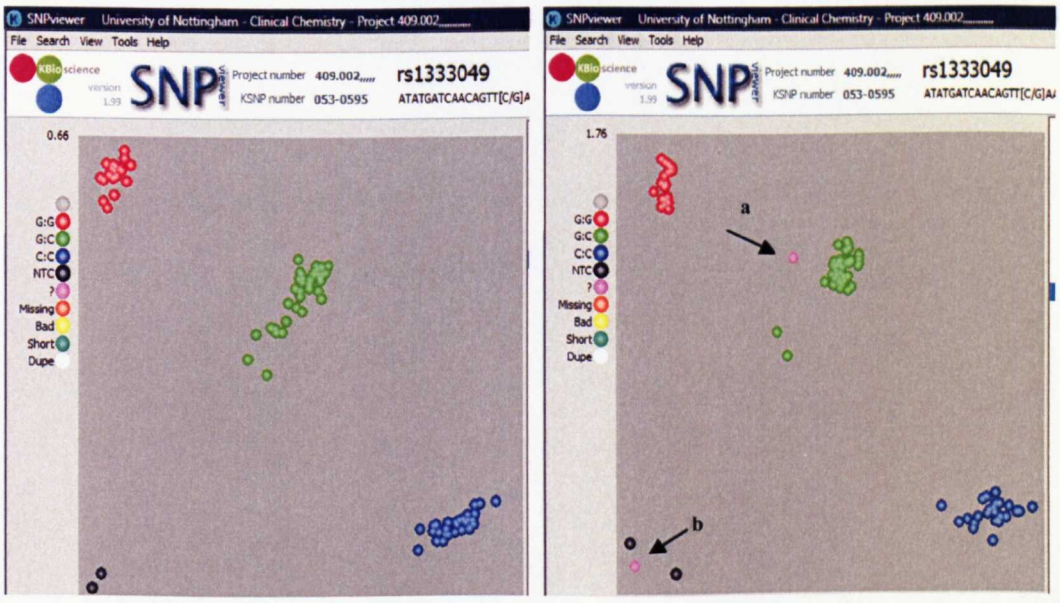
Early stages of PCR, showing the DNA template (black), one of the forward allele specific primers with its unique unlabelled tail (green) and the common reverse primer (yellow). The FAM labelled oligo (red) binds its complementary quencher oligo (Blue) early during the PCR reaction. B) Late in the PCR reaction the allele specific forward primer with non labelled tail, after being incorporated into the PCR product, binds to the FAM labelled oligo and gives rise to fluorescent energy.



The data can be viewed graphically by software (KlusterKaller) in which the FAM and Cal Fluor 560 data are plotted on the x and y axes, respectively. Data were supplied by KBioscience as a spread sheet with the genotype already specified, which were double checked on receipt using the SNPviewer (http://www.kbiosciences.co.uk/download/snpviewer) according to sample clusters (Figure 2-11). An average of 4 samples per plate was re-called as a result of this manual checking procedure.

Figure 2-11: SNPviewer of samples genotyped for SNP rs1333049

The left hand panel shows one of the plates in which all samples were successfully genotyped. The right hand panel shows 2 ambiguous samples; sample (a) was heterozygous GC for SNP rs1333049, as it was close to the heterozygous cluster, sample (b) could not be re-called as it was close to the NTC samples.



2.4.5.2 Mutagenically separated PCR (MS-PCR)

MS-PCR is a single step PCR based genotyping method, which relies on the instability of the primer when annealing with the DNA template if there is a mismatch at the 3'terminal base. To increase the specificity of primer annealing another mismatch is created at the third base from the 3'end of the primer (Rust et al., 1993).

There are 2 methods of primer design for MS-PCR, the three-primer and four-primer methods. The three primer method was used for genotyping the TCF7L2-rs7903146, C>T polymorphism. By using allele specific primers of different length, two PCR products of differing size can be generated in the same reaction. This difference in length is achieved by the use of a non specific tail at the 5'end of one primer (Figure 2-12). The four-primer method was used for genotyping of the PPAR-γ rs18012182, C>G polymorphism. In the four-primer method, the primer locations were selected to ensure that the PCR products generated by the 2 allele specific primers and the 2 external primers are of differing size (Figure 2-13). The PCR products can be viewed on agarose gel directly without the need for further manipulation.

The sequence of the primers used for genotyping of SNPs rs7903146 and rs18012182 are presented in

.

- Each primer set was designed using Primer3 software (http://www.primer3plus.com/cgi-bin/dev/primer3plus.cgi).
- The specificity of each primer pair was tested using both the blast function embedded in primer 3 (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and

the blat function in the UCSC genome browser (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start).

- Specific software was used to test whether primers were located over other polymorphic sites which could interfere with primer annealing using diagnostic SNPcheck from Manchester (https://ngrl.manchester.ac.uk/SNPCheck V2/snpcheck.htm).
- All primers were supplied by Eurogentec S.A., Belgium.
- Primers were kept at a stock concentration of 100μmol/μl, and working aliquots of 10μmol/μl were prepared to avoid contamination.

Table 2-4: Primers used for MS-PCR

Two SNPs were genotyped using MS-PCR. A three-primer method was used for genotyping TCF7L2 rs7903146 and four-primer method for genotyping PPAR-γ rs18012182. The SNPs and the third base at the 3'end mismatch of the primer are shown in bold. The non specific tail at the 5'end the T allele primer is underlined.

Primer	Sequence
TCF7L2 rs7903146 (C>T)	
Common forward primer	5'-AAAGGGAGAAAGCAGGATTG-3'
C allele reverse primer	5'-GCCTCATACGGCAATTAAATTATCAG-3'
T allele reverse primer	5'-<u>CACTAAGGACGCAATGATTGT</u>GCCTCATA CGGCAATTAAATTATG**A**A-3'
PPAR-γ rs18012182(C>G)	
Common forward primer	5'- GTACAGTTCACGCCCCTCAC-3'
Common reverse primer	5'-TGCTTAGCTCGTTGCTATCG-3'
C allele specific reverse primer	5'-GTGAAGGAATCGCTTTCAG**G**-3'
G allele specific forward primer	5'-TGGGAGATTCTCCTATTGTC**G**-3'

Figure 2-12: MS-PCR, three-primer method

Three-primer MS-PCR applied to the C/T polymorphism at rs7903146 (in red bold) using a common forward primer (black arrow) and 2 allele specific reverse primers. The terminal 3' base is allele specific, and the 3rd base at the 3' end (underlined) is mismatched to increase the specificity of annealing. To discriminate between PCR products a 20 bp tail of scrambled sequence is added at the 5' end of the T allele reverse primer (red arrow), so that the C allele will yield PCR product of 212 bp, while the T allele will yield PCR product of 232 bp.

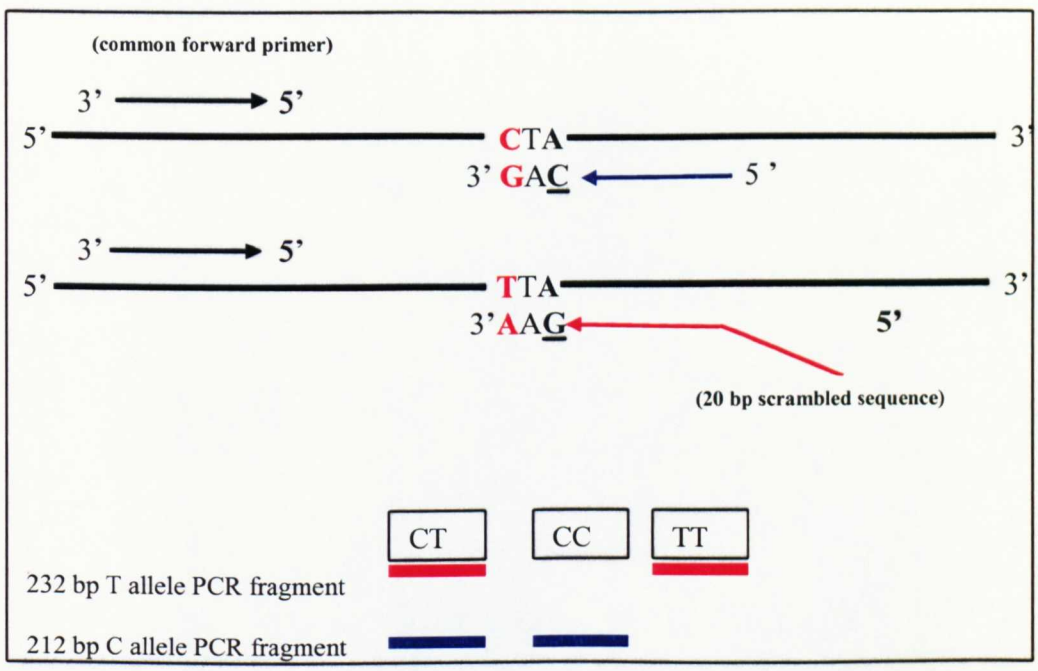
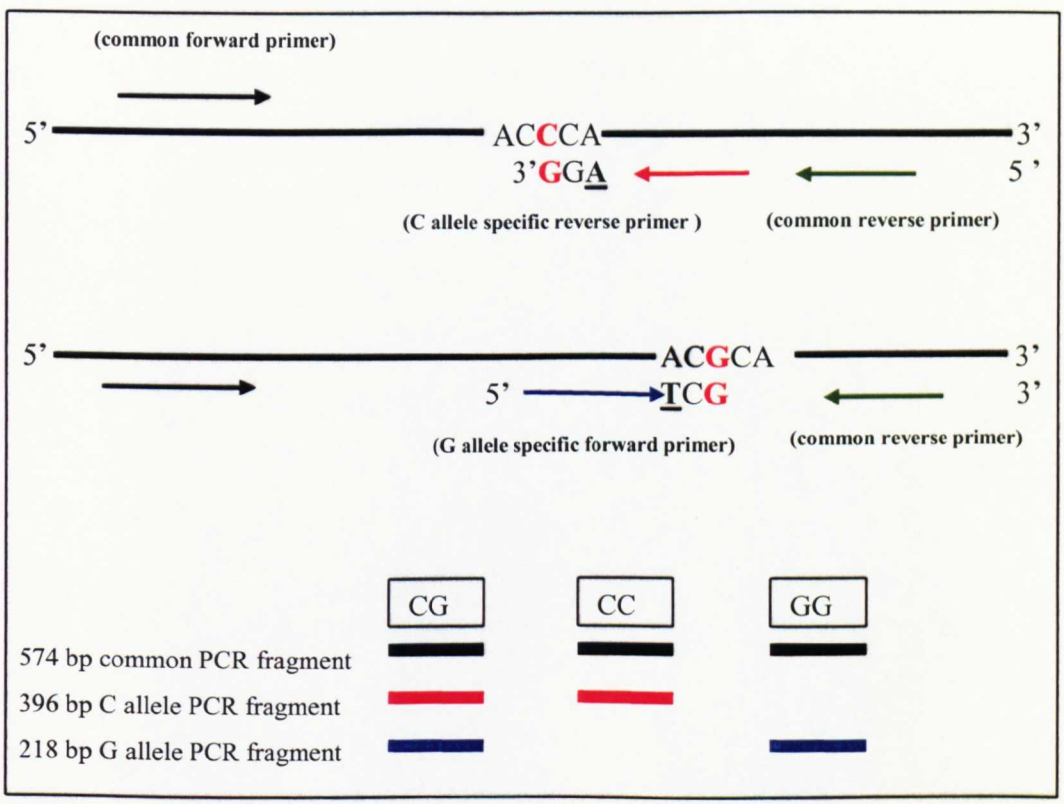


Figure 2-13: MS-PCR, four-primer method

Four-primer MS-PCR applied to the C/G of polymorphism rs12812182 (in red bold) using a forward (black arrow) and reverse primers (green arrow) and 2 allele specific primers, a C allele specific reverse primer (red arrow) and a G allele specific forward primer (blue arrow). The terminal 3' base is allele specific, and further mismatch is introduced at the 3rd base at the 3' end (underlined) to improve the specificity of annealing. The primer locations were selected to ensure that the PCR products generated by the 2 allele specific primers and the external primers are of different sizes.



2.4.5.3 PCR reaction and condition

PCR reactions were optimized for annealing temperature and primer concentration. The final PCR reagent concentrations and reaction conditions are shown in Table 2-5 and Table 2-6.

Table 2-5: MS-PCR reaction for SNP rs7903146, and rs18012182

	SNP rs79023146	SNP rs18012182
Buffer	1X PCR buffer (750 mM Tris-HCl, 200 mM $(NH_4)_2SO_4$, 0.1% (v/v) Tween 20)	1x Paq5000®Hotstart PCR master mix(optimized PCR buffer, dNTPs, Mg Cl_2, DNA polymerase)
dNTPs	200 μM	1mM
Mg Cl_2	1.5mM	2.25mM
Other additives	1x Q-solution	5% DMSO
DNA polymerase	2U Taq polymerase	Paq5000®Hotstart DNA polymerase
DNA	200 ng genomic DNA	200 ng genomic DNA
Forward primer(s)	1 μM TCF-F	0.5 μM common forward primer 0.5 μM G allele forward primer
Reverse primers	1 μM TCF-C primer 0.3 μM TCF-T primer	0.5 μM common reverse primer 0.25 μM C allele reverse primer
PCR grade water	To 30μl total volume	To 30μl total volume

Table 2-6: PCR and electrophoresis conditions for SNP rs7903146 and rs1801282

	SNP rs79023146	SNP rs12812182
	PCR conditions	
Initial denaturation	94°C /5min	95°C/2 min
Number of cycles	35	30
Denaturation	94°C/ 40s	95°C/ 20s
Annealing	61°C /1min	56°C/20s
Extension	72°C /1min	72°C/ 30s.
Final extension step	72°C/10min	72°C/10min.
Thermal cycler	MJ Research PTC 225 Thermal Cycler, GMI Inc, USA	Veriti™ 96-Well thermal cycler/Applied Biosystems
	Electrophoresis conditions	
Agarose gel	3.5% in 1xTAE with 5µl ethidium bromide	1% in 1xTAE with 5µl of ethidium bromide
Electrophoresis	60V/ 45min	80V/25min
Size of PCR product	212bp for c allele 232bp for T allele	574 bp common PCR fragment 396bp for C allele 218bp for G allele

2.4.5.4 Gel electrophoresis

MS-PCR products were resolved by electrophoresis on agarose gel (Eurogentec) prepared in 1xTAE (Tris acetate EDTA) buffer and stained with ethidium bromide 5µl/ 25 ml. The 1xTAE was diluted from a 50xTAE stock solution (242gTris base, 57.1 ml glacial acetic acid, 100 ml 0.5M EDTA (Ethylenediaminetetraacetic acid) pH8.0) per litre). 10µl of PCR product were mixed with 2µl of 6x DNA loading dye (Fermentas). A 100 bp DNA ladder was used as a size marker (Fermentas, Gene Ruler™). An image of the gel was captured on a UV transilluminator (AlphaDigiDoc™).

2.4.5.5 DNA sequencing

PCR products were cleaned using ExoSAP-IT®, which contains shrimp alkaline phosphatase to remove un-used dNTPs and an exonuclease to digest the primers that can interfere with the sequencing reaction. 2µl ExoSAP were added to 5µl of the PCR reaction and incubated at 37°C for 15 min. ExoSAP was inactivated by heating to 80°C for 15 min.

The sequencing reaction was carried out in a 10µl reaction in which 3µl of cleaned PCR product was mixed with 2.5X Big Dye 3.1 reaction mix (containing the fluorescent labelled ddNTPs and polymerase enzyme) 5x big dye terminator buffer (Applied Biosystems), and 5pmol of the common forward primer. Cycling conditions were as follows, 96 °C for 1 min followed by 25 cycles of 96°C for 30 seconds, 50°C for 15s, and 60°C for 4 min for the SNP rs7903146. The cycling condition for SNP rs1801282 was 1 min at 97°C, followed by 25 cycles of 10s at 96°C, 5s at 50°C and 4 min at 60°C. The difference in the cycling conditions was related to the use of different PCR blocks. The samples were then held at 10°C until collection.

Samples were centrifuged at 5000 rpm for 3 min through PERFORMA® Spin Columns (Edge Biosystems).The eluent containing the DNA was then transferred into 0.5 ml tubes and dehydrated on a heat block at 90°C. Prior to sequencing on an ABI 3130 genetic analyzer (Applied Biosystems), the samples were treated with 20µl template suppression reagent, denatured at 95°C degrees for 2 min, chilled on ice, then loaded for capillary electrophoresis. Sequencing results were viewed by Chromas Lite V2.01 (http://www.technelysium.com.au/

chromas_lite.html), and sequence alignment was carried out by ClustalW web based software (http:/www.ebi.ac.uk/Tools /clustalw2/index.html).

2.4.6 Genotyping quality control

Several procedures were put in place to ensure that the genotyping was carried out correctly.

2.4.6.1 Quality control for Kbioscience genotyping

- 5% of the samples were included as inter and intra plate duplicates.
- Two NTC samples per plate were included.
- Case and control samples were randomly dispensed into the plates.
- Any plate or any SNP with <90% call rate was repeated.
- Allele frequency and concordance with Hardy Weinberg equilibrium (HWE) were calculated for each SNP on individual plates by Kbiosciences.
- Each plate was checked on the SNPviewer to assure accurate calling of the genotype. It was possible using this strategy to call an additional 4 samples in some plates (Figure 2-11).
- 200 and 300 samples for SNP rs7903146 and rs1801282 respectively were genotyped in house by MS-PCR and checked for concordance with the Kbiosciences results.

2.4.6.2 Quality control for in house genotyping

- A DNA-free negative control was included in each PCR run to detect contamination of the reagents.
- Heterozygous samples confirmed by DNA sequencing, were included in each PCR run as positive controls.

- Genotypes were assigned by two independent investigators, who were blinded to which samples were cases and which were controls. Samples assigned discordant genotypes were re-genotyped.
- 6 and 13 samples for rs7903146 and rs1801282 respectively representing all three genotypes were selected for sequencing as a quality assurance measure.
- The whole sample set and the controls were tested for Hardy Weinberg equilibrium (HWE) using the web based Hardy Weinberg equilibrium calculator (http://www.oege.org/software /hwe-mr-calc.shtml).

2.4.7 Statistical analysis

Genotype frequency and genotype relative risk were analyzed using an additive model of inheritance after scoring the genotypes as 1, 2, 3 for major allele homozygous, heterozygous, and minor allele homozygous respectively, using binary logistic regression (SPSS version 16). The UNPHASED program was used for the analysis of allelic OR and haplotype analysis. This program uses the maximum likelihood ratio in a log linear model for the analysis of allelic and haplotype effects. SPSS was used for phenotypic and quantitative analysis. A P value of ≤ 0.01 was selected as the threshold for statistical significance, after adjusting for testing multiple hypotheses. Bonferroni correction to adjust for each of the SNPs tested in the 2 conditions studied (preeclampsia and FGR) was considered over conservative due to lack of independence between SNPs and the claimed relationship between preeclampsia and FGR.

2.5 Results

2.5.1 Quality control and population genetics results

The genotype call rate for the loci investigated was > 98%, (Table 2-7). Within the Kbiosciences genotyped samples, duplicates were 100% concordant. Initially, 15 (7.5%) and 6 (2%) of the samples genotyped by MS-PCR for SNPs rs7903146 (Figure 2-14) and rs1801282 (Figure 2-15) respectively, were not concordant with the Kbiosciences mainly due to poor visibility on the agarose gel, as illustrated in figure 2-14, lanes A and E. These samples were subjected to regenotyping and sequencing (Figure 2-16). Two and one samples for SNPs rs7903146 and rs1801282 respectively, remained discordant and were excluded from the study.

Figure 2-14: Electophoresis banding pattern for SNP rs7903146

MS-PCR for rs79023146 C>T polymorphism. B, C and D are homozygous CC heterozygous CT, and homozygous TT samples respectively. A and E are samples that were difficult to recall. F is a sequenced heterozygous sample used as a positive control, showing the T allele at 232-bp and the C allele at 212-bp. G is a DNA free negative control.

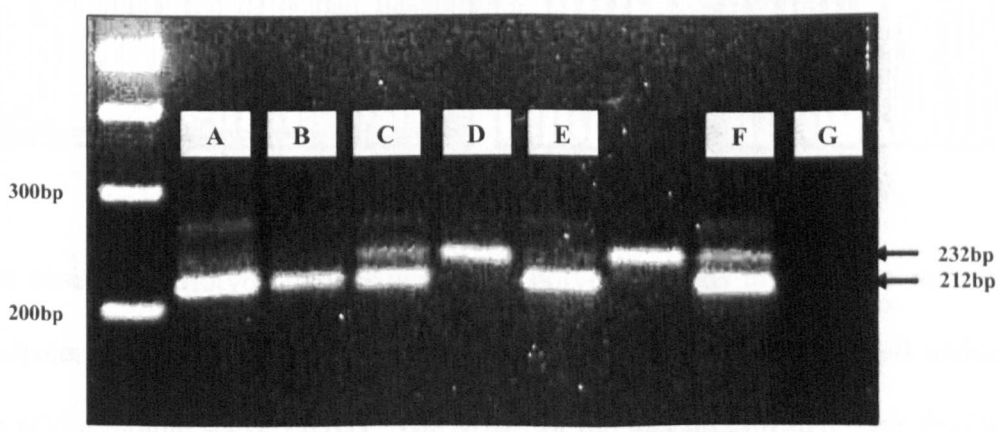


Figure 2-15: Electrophoresis banding pattern for SNP rs1801282

MS-PCR for rs1801282 C>G polymorphism. A, B, and C are homozygous CC, heterozygous CG and homozygous GG samples respectively. D is a sequenced heterozygous sample used as a positive control showing the 574bp common PCR product, 396bp C allele product and the 218bp G allele product. E is a DNA free negative control.

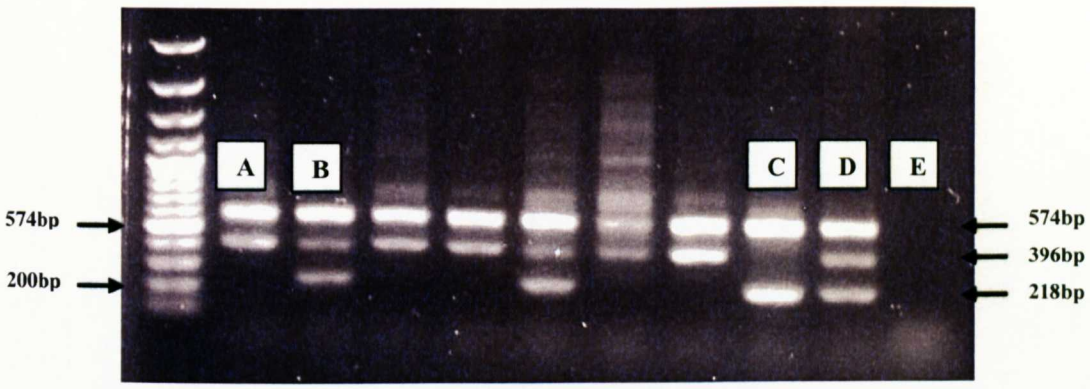
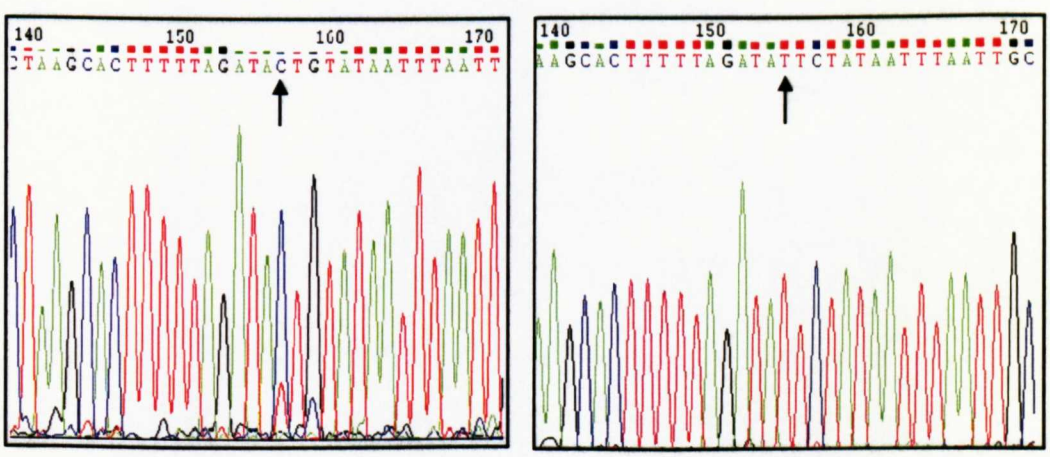


Figure 2-16: Electropherogram showing CC and TT genotypes for SNP rs7903146



The observed genotype frequencies of the polymorphisms investigated were in agreement with HWE within individual plates, in the control group and within the whole sample set ($P > 0.05$). The MAF for the loci investigated were similar to that reported in the Hap Map project for the European population, and with that recorded in GWAS studies (Table 2-7).

Table 2-7: Minor allele frequency and quality control data for study variants

Chr = Chromosomal location based on Hap Map release#24. * Major allele confers risk for T2D. MAF: minor allele frequency ** Hardy Weinberg Equilibrium calculated for the entire sample.

Gene	SNP	Location	Major/ Minor allele	MAF	Success rate (%)	Hardy Weinberg** (P value)
TCF7L2	rs7903146	chr10:114748339	C/T	0.29	98.0	0.06
	rs7924080	chr10:114777002	T/C	0.47	98.6	0.45
	rs12255372	chr10:114798892	G/T	0.29	99.5	0.18
FTO	rs17817449	chr16:52370868	T/G	0.41	98.6	0.65
PPAR-γ	rs1801282	chr3:12368125	C/G*	0.12	98.6	0.22
	rs3856806	chr3:12450557	C/T*	0.13	98.8	0.51
	rs10865710	chr3:12328198	C/G	0.24	98.8	0.64
CDKN2B-	rs10811661	chr9:22124094	T/C*	0.17	98.6	0.05
AS1	rs1333049	chr9:22115503	G/C	0.49	99.0	0.51
KCNJ11	rs5215	chr11:17365206	T/C	0.35	98.4	0.32
	rs5219	chr11:17366148	C/T	0.35	98.2	0.27

2.5.2 Preeclampsia cases and controls: Phenotypic Data

Tables 2-8 and 2-9 present details of phenotypic data for the preeclampsia case control study. There were no significant differences in the maternal age between preeclampsia cases and controls. Women with preeclampsia had significantly lower BMI (Table 2-8), reflecting the inclusion of obesity as a risk factor for preeclampsia in the VIP study. By definition systolic and diastolic blood pressure measures were higher in preeclampsia cases compared to controls (Table 2-8). Unsurprisingly, preeclamptic women delivered at earlier gestational ages and gave birth to babies with lower birth weight compared to controls (Table 2-8).

Table 2-8: Phenotypic data for preeclampsia and healthy pregnancy

IQR = interquartile range. P values were calculated using the Mann Whitney U test as the data were not normally distributed.

	Controls (N=673)		Preeclampsia (N=448)		P value
	Median	IQR	Median	IQR	
Maternal age (years)	31	26-35	32	27-35	0.46
BMI(kg/m2)	30	24-35	28	24-32	<0.001
Maximum systolic blood pressure (mmHg)	120	110-130	147	120-162	<0.001
Maximum diastolic blood pressure (mmHg)	70	60-80	98	72-106	<0.001
Gestational age at delivery (days)	273	265-283	266	252-275	<0.001
Fetal Weight (gm)	3350	2989-3730	3040	2299-3540	<0.001
Birth weight centile	47	25-77	38	5-73	<0.001

More primiparous women were present in the preeclampsia group compared to controls, with primiparity conferring risk for preeclampsia (OR = 2.15, 95%CI = 1.60-2.87, P<0.001). Data on smoking habits was available for 659 (98%) controls and 433 (97%) preeclampsia cases. Preeclamptic mothers were more likely to be non-smokers and less likely to be previous or current smokers, although this was not statistically significant. The proportion of offspring of each sex did not differ between cases and controls (Table 2-9).

Table 2-9: Parity, smoking habit and baby's sex in preeclampsia and healthy pregnancy

P values were calculated using Chi-squared test. * Data on smoking habits was available for 659 (98%) controls and 433 (97%) preeclampsia cases, ** fetal sex included for singleton pregnancies only.

		Controls N (%)	Preeclampsia N (%)	P value
Parity	Primiparous	459 (68.2)	368 (82.1)	<0.001
Smoking habit*	Never smoker	365 (55.4)	260 (60.0)	0.306
	Stopped smoking	194 (29.4)	116 (26.8)	
	Current smoker	100 (15.2)	57 (13.2)	
Baby's sex**	Male	336 (52.7)	227 (51.7)	0.737
	Female	301 (47.3)	212 (48.3)	

343 preeclampsia cases and 574 controls were volunteers with risk factors for preeclampsia, recruited to the VIP study. Within this subgroup the rate of chronic hypertension, diabetes, abnormal uterine artery Doppler, chronic renal disease and anti-phospholipid syndrome were similar in cases and controls and similar to the original VIP study. However, the rate of multiple pregnancies and women with BMI >30 in the first pregnancy were higher in the original VIP trial and in controls compared to preeclampsia cases. The distribution of risk factors in preeclampsia cases, controls, and in the original VIP trial is presented in Table 2-10. 240 women with a history of preeclampsia in the pregnancy preceding the VIP pregnancy were considered cases for the purpose of the genetic study (Table 2-10).

Table 2-10: Frequency of risk factors in preeclampsia and control groups recruited from the VIP collection

Data are presented as Number (%). P values were calculated using the Chi-squared test for the comparison between cases and controls. Previous Preeclampsia, HELLP or eclampsia were considered as cases for the purpose of genetic studies.

Risk factor	VIP study N=1988	Genetic study		P value
		Controls N= 574	Preeclampsia N=343	
Previous Preeclampsia, HELLP or eclampsia	548 (23)	-	240 (70.0)	<0.001
Chronic hypertension	869 (36)	185 (32)	128 (37)	0.12
Diabetes	199 (8)	56 (10)	23 (7)	0.11
Abnormal uterine artery Doppler	80 (3)	10 (2)	7 (2)	0.75
Multiple pregnancy	503 (16)	103 (18)	21 (6)	<0.001
Chronic renal disease	47 (2)	7 (1)	8 (2)	0.20
Anti-phospholipid syndrome	52 (2)	12 (2)	8 (2)	0.80
BMI > 30(kg/m^2) in primiparous women	804 (33)	274 (48)	40 (12)	<0.001
Multiple risk factors	503 (21)	71 (12)	114 (33)	<0.001

43 women had preeclampsia in both the pregnancy preceding and during the VIP study implying a 17.9% recurrence rate. Amongst women with chronic hypertension, 128 (40.6%) women developed preeclampsia. Preeclampsia developed in 23 (29.1%) and 8 (53.3%) diabetic women and in women with chronic renal failure respectively. Seven (41.8%) and 8 (40.0%) women with abnormal uterine Doppler and antiphospholipid syndrome respectively developed preeclampsia. The rate of preeclampsia in women with risk factors in this study and in the original VIP study is reported in Table 2-11.

Table 2-11: Rate of preeclampsia in women with risk factors recruited from the VIP collection

VIP study recruited 1988 white western European women, from which 917 were included in the preeclampsia case control genetic study. Data are presented as Number (%).

	VIP study N (%)	Preeclampsia genetic study N (%)
Previous Preeclampsia, HELLP or eclampsia	124 (23)	43 (18)
Chronic hypertension	192 (23)	128 (41)
Diabetes	34 (16)	23 (29)
Abnormal uterine artery Doppler	16 (25)	7 (42)
Multiple pregnancy	41 (11)	21 (17)
Chronic renal disease	15 (32)	8 (53)
Anti-phospholipid syndrome	6 (23)	8 (40)
BMI > 30(kg/m^2) in primiparous women	87 (12)	40 (13)
Multiple risk factors	127 (26)	114 (33)

2.5.3 Preeclampsia case control genetic association study

The G allele of the marker rs17817449 on the FTO gene was found to be potentially protective against preeclampsia (OR 0.84, 95% CI 0.71-1.00, P value 0.05) an effect which was abolished after adjustment for parity, BMI, smoking habit and risk factors including chronic hypertension, diabetic status, chronic renal disease, abnormal uterine artery Doppler, and multiple pregnancy (OR 0.89, 95% CI 0.74-1.09, P value 0.27), (Table 2-12).

There were no differences in the genotype, allele or haplotype frequencies between the preeclampsia and control groups in TCF7L2, PPAR-γ, CDKN2B-AS1 or KCNJ11 variants. The 95% CI for the unadjusted and adjusted OR all span 1 and P value was>0.05 (Table 2-12 and Table 2-13).

Table 2-12: Genotype and allele distributions of T2D alleles and corresponding odds ratios for preeclampsia

Genotype OR was calculated using binary logistic regression, using SPSS v.16. * adjusted for parity, smoking habit, BMI and the risk factors including chronic hypertension, diabetic status, chronic renal disease, abnormal uterine artery Doppler, and multiple pregnancies. ** when the genotype frequency was < 5% the heterozygotes and the rare homozygotes were pooled for analysis. Allelic OR was calculated using UNPHASED software release 3.1.5.

Gene	SNP rs number Genotype	Controls N= 673	Cases N=448	Genotype unadjusted OR (95%CI) P value	Genotype adjusted OR* (95%CI) P value	Unadjusted allelic OR (95% CI) P value
TCF7L2	rs7903146					
	CC	53.2	50.2	1.05	1.08	**1.05**
	CT	37.3	41.0	(0.88-1.27)	(0.88-1.31)	**(0.87-1.27)**
	TT	9.5	8.8	0.58	0.44	**0.60**
	T	**28.3**	**29.3**			
	rs7924080					
	TT	30.8	28.3	1.01	1.04	**1.01**
	TC	47.6	51.8	(0.85-1.20)	(0.87-1.25)	**(0.85-1.20)**
	CC	21.6	19.9	0.89	0.68	**0.88**
	C	**45.4**	**45.8**			
	rs12255372					
	GG	51.0	54.3	0.87	0.89	**0.87**
	GT	39.1	37.8	(0.72-1.05)	(0.73-1.091	**(0.72-1.05)**
	TT	9.9	7.9	0.157	0.27	**0.14**
	T	**29.3**	**26.4**			
FTO	rs17817449					
	TT	30.6	34.5	0.84	0.89	**0.84**
	TG	49.1	48.7	(0.71-1.00)	(0.74-1.09)	**(0.71-1.00)**
	GG	20.3	16.8	0.05	0.27	**0.05**
	G	**43.7**	**39.6**			
PPAR-γ	rs1801282					
	CC	78.2	79.0	1.15	1.17	**1.11**
	CG	20.4	19.3	(0.87-1.53)	(0.86-1.59)	**(0.86-1.43)**
	GG**	1.4	1.7	0.33	0.30	**0.44**
	G	**13.1**	**12.0**			
	rs3856806					
	CC	75.8	77.8	1.09	1.14	**1.12**
	CT	23.2	19.7	(0.82-1.44)	(0.8451.54)	**(0.87-1.43)**
	TT**	1.1	2.5	0.555	0.38	**0.39**
	T	**12.7**	**14.0**			
	rs10865710					
	CC	57.8	57.3	1.05	0.965	**1.06**
	CG	37.6	36.4	(0.83-1.34)	(0.743-1.25)	**(0.87-1.28)**
	GG	4.6	6.3	0 .67	0.79	**0.58**
	G	**24.2**	**25.2**			

Gene	SNP rs number Genotype	Controls N= 673	Cases N=448	Genotype unadjusted OR (95%CI) P value	Genotype adjusted OR* (95%CI) P value	Unadjusted allelic OR (95% CI) P value
CDKN2B-AS1	rs10811661					
	TT	70.8	64.6	1.16	1.12	**1.08**
	TC	25.4	30.8	(0.89-1.50)	(0.84-1.50)	**(0.86-1.36)**
	CC**	3.7	4.6	0.28	0.45	**0.52**
	C	**16.1**	**17.2**			
	RS1333049					
	GG	26.1	30.1	0.92	0.912	**0.92**
	GC	48.6	47.3	(0.77-1.08)	(0.76-1.10)	**(0.77-1.08)**
	CC	25.4	22.6	0.30	0.32	**0.29**
	C	**50.4**	**48.1**			
KCNJ11	Rs5215					
	TT	42.9	42.3	1.09	1.10	**1.08**
	TC	44.3	43.6	(0.81-1.29)	(0.91-1.33)	**(0.91-1.29)**
	CC	12.8	14.1	0.36	0.32	**0.38**
	C	**35.2**	**37.07**			
	Rs5219					
	CC	42.9	41.8	1.09	1.10	**1.09**
	CT	44.1	44.4	(0.91-1.29)	(0.91-1.33)	**(0.91-1.29)**
	TT	13.0	13.8	0.34	0.31	**0.36**
	T	**35.2**	**37.1**			

Table 2-13: Haplotype frequency for T2D susceptibility genes in preeclampsia cases and controls

Haplotype with frequency <5% were pooled as rare haplotypes, P values were calculated using likelihood ratio Chi-squared for the common haplotypes only. UNPHASED V3.1.5 was used for this analysis.

TCF7L2				
Rs7903146 (C/T)	Rs7924080 (T/C)	Rs12255372 (G/T)	**Cases**	**Controls**
C	T	G	0.52	0.52
C	C	G	0.17	0.17
T	C	T	0.24	0.23
Rare haplotypes			0.07	0.08
P value			**0.99**	
PPAR-γ				
Rs1801282 (C/G)	Rs3856806 (C/T)	Rs10856710 (C/G)	**Cases**	**Controls**
C	C	C	0.71	0.73
C	C	G	0.11	0.11
G	T	G	0.09	0.08
Rare haplotypes			0.09	0.08
P value			**0.77**	
CDKN2B-AS1				
Rs10811661 (T/C)	Rs1333049(G/C)	-	**Cases**	**Controls**
T	G	-	0.42	0.42
T	C	-	0.41	0.42
C	G	-	0.09	0.09
C	C	-	0.08	0.07
P value			**0.63**	
KCNJ11				
Rs5215 (T/C)	Rs5219 (C/T)	-	**Cases**	**Controls**
T	C	-	0.63	0.65
C	T	-	0.37	0.35
Rare haplotypes		-	0.012	0.015
P value		-	**0.39**	

2.5.4 Fetal growth restriction (FGR) phenotypic data

There were no significant differences between FGR cases and controls regarding maternal age. Mothers with FGR cases had lower BMI, (Table 2-14). Maternal BMI was positively correlated with birth weight in FGR cases (r = 0.15) and controls (r= 0.19), P<0.001. Systolic and diastolic blood pressure measures were higher in FGR cases compared to controls (Table 2-14). FGR mothers were delivered at earlier gestational ages. By definition birth weight and birth weight centile were lower in FGR cases compared to controls (Table 2-14).

There was no variation in the frequency of primiparity in FGR cases and controls. Data on smoking habits was available for 540 (95%) of the controls and 239 (98%) of the FGR cases. Smoking was associated with a higher risk of FGR (OR=1.2, 95% CI 0.99-1.45, P=0.05). Male and female babies were equally distributed in FGR cases and controls. Female babies were lighter than male babies by about 140 gm (P=0.001) in FGR controls but did not differ within the FGR cases. These data are summarized in Table 2-15.

Table 2-14: Phenotypic data for FGR and healthy pregnancy

IQR = interquartile range. P values were calculated using the Mann Whitney U test as the data were not normally distributed.

	Controls (N=570)		FGR (N=243)		P value
	Median	IQR	Median	IQR	
Maternal age (years)	30	26-34	30	24 -34	0.11
BMI(kg/m2)	31	25-35	25	22-30	<0.001
Maximum systolic blood pressure (mm Hg)	120	110-130	124	110-160	<0.001
Maximum diastolic blood pressure (mm Hg)	70	64-80	74	60-104	0.001
Gestational age at delivery (days)	275.5	267-285	259	231-273	<0.001
Fetal WT (gm)	3430.0	3176-3803	2145	1432.5-2492	<0.001
Birth weight centile	52	28-80	1	0-2	<0.001

Table 2-15: Parity, smoking habit and baby's sex in FGR and healthy pregnancy

P values were calculated using Chi-squared test. * Data on smoking habits was available for 540 (95%) of the controls and 239 (98%) of the FGR cases.

		Controls N (%)	FGR N (%)	P value
Parity	Primiparous	399 (70.0)	163 (67.1)	0.41
Smoking habit*	Never smoker	311 (55.7)	134 (56.1)	<0.001
	Stopped smoking	161 (28.9)	40 (16.7)	
	Current smoker	68 (15.4)	65 (27.2)	
Fetal sex	Male	310 (54.4)	121 (49.8)	0.38
	Female	259 (45.4)	122 (50.2)	

110 FGR cases and 471 controls were volunteers with risk factors for preeclampsia, recruited to the VIP study. Within the subgroup used in our study the rates of previous preeclampsia, HELLP syndrome, and eclampsia, and chronic renal disease were higher in FGR cases than controls. Predictably, the proportion of women with BMI >30 in the first pregnancy was higher in controls compared to FGR cases. The frequency of chronic hypertension, diabetes, anti-phospholipid syndrome and abnormal uterine artery Doppler were similar in FGR cases and control. Distributions of risk factors in preeclampsia cases, controls, and in the original VIP trial are presented in Table 2-16.

Table 2-16: Frequency of risk factors in FGR and control groups recruited from VIP collection

Data are presented as Number (%). P values were calculated using the Chi-squared test for the comparison between cases and controls. Previous Preeclampsia, HELLP or eclampsia were excluded from the control group.

	VIP study N=1988	Controls N=471	FGR N=110	P value
Previous Preeclampsia, HELLP or eclampsia	548 (23)	0	36 (32)	<0.001
Chronic hypertension	869 (36)	178 (38)	50 (45)	0.14
Diabetes	199 (8)	54 (12)	9 (8)	0.32
Abnormal uterine artery Doppler	80 (3)	10 (2)	5 (4)	0.15
Chronic renal disease	47 (2)	7 (2)	7 (6)	0.003
Anti-phospholipid syndrome	52 (2)	12 (3)	3 (2)	0.92
BMI > 30 (kg/m^2) in primiparous women	804 (33)	267 (58)	36 (32)	< 0.001
Multiple risk factors	503 (21)	57 (12)	32 (29)	< 0.001

32 women who had preeclampsia in the pregnancy preceding the VIP study developed FGR during the course of the VIP trial. Within the chronically hypertensive women, 50 (22.0%) developed FGR. FGR developed in 9 (14%) and 7 (50%) diabetic women and in women with chronic renal disease respectively. 5 (33%) and 3 (20.0%) women with abnormal uterine Doppler and women with antiphospholipid syndrome respectively had pregnancies complicated by FGR. The rate of FGR in women with risk factors in this study and in the original VIP study is reported in Table 2-17.

Table 2-17: Rate of FGR in women with risk factors within the VIP collection

Data are presented as Number (%).

	VIP study N (%)	FGR N (%)
Previous Preeclampsia, HELLP or eclampsia	124 (23)	36 (33)
Chronic hypertension	192 (23)	50 (22)
Diabetes	34 (35)	9 (14)
Abnormal uterine artery Doppler	16 (25)	5 (33)
Chronic renal disease	15 (32)	7 (50)
Anti-phospholipid syndrome	6 (23)	3 (20)
BMI > 30(kg/m^2) in primiparous women	87 (12)	36 (12)
Multiple risk factors	127 (26)	32 (29)

2.5.5 Fetal growth restriction (FGR) case control genetic association study

The C allele of rs10811661 of the CDKN2B-AS1 gene showed suggestive association with FGR (OR=1.33, 95% CI 0.97-1.84, P=0.08), (Table 2-18). This effect was lessened after adjustment for smoking habit and risk factors including chronic hypertension, diabetic status, chronic renal disease, abnormal uterine artery Doppler, (OR=1.26, 95% CI 0.91-1.75, P=0.17), (Table 2-18), (note that adjustment for parity, BMI and fetal sex is included in the calculation of birth weight centile). SNP rs1333049 alone showed no effect on FGR risk (OR=0.88, 95%CI (0.71-1.08), P=0.22), (Table 2-18). The CG haplotype of the CDKN2B-AS1 gene formed of the diabetes protective allele of SNP rs10811661 and the coronary artery disease protective allele of SNP rs1333049 was associated with a higher risk for FGR (P 0.005), (Table 2-19). When compared with the TG haplotype (formed of the major alleles), the CG haplotype had an OR of 1.98, (95% CI 1.33-2.95, P=0.0008). The observed haplotype effect suggests interaction between the 2 loci under investigation as the 2 variants of the CDKN2B-AS1 gene lie in different LD blocks ($r^2 = 0.0003$), (Figure 2-6).

Variants of TCF7L2, FTO, PPAR-γ, and KCNJ11 showed no association of genotype, allele or haplotype frequencies with FGR before or after adjustment for possible confounding factors (Table 2-18 and Table 2-19)

Table 2-18: Genotype and allele distributions of T2D alleles and corresponding odds ratios for fetal growth restriction (FGR)

Genotype OR was calculated using binary logistic regression (SPSS v.16). * Adjusted smoking habit, and risk factors including chronic hypertension, diabetic status, chronic renal disease, and abnormal uterine artery Doppler. ** When the genotype frequency was < 5% heterozygotes and rare homozygotes were pooled for analysis. Allelic OR was calculated using UNPHASED release 3.1.5.

	SNP rs number Genotype	Controls N=570	FGR cases N=243	Genotype unadjusted OR (95%CI) P value	Genotype adjusted OR* (95%CI) P value	Unadjusted allelic OR (95% CI) P value
	rs7903146					
	CC	53.1	51.3	0.99	0.99	**0.99**
	CT	37.3	41.0	(0.79-1.26)	(0.79-1.27)	**(0.79-1.27)**
	TT	9.6	7.7	0.98	0.99	**0.98**
	T	**28.2**	**28.3**			
	rs7924080					
	TT	30.8	23.5	1.16	1.15	**1.17**
TCF7L2	TC	47.6	54.6	(0.94-1.44)	(0.93-1.43)	**(0.94-1.44)**
	CC	21.6	21.8	0.17	0.21	**0.16**
	C	**45.2**	**49.7**			
	rs12255372					
	GG	51.0	54.4	0.88	0.88	**0.88**
	GT	39.2	37.8	(0.69-1.11)	(0.69-1.11)	**(0.69-1.11)**
	TT	9.9	7.9	0.29	0.28	**0.27**
	T	**29.5**	**26.8**			
	rs17817449					
	TT	30.6	34.5	0.86	0.85	**0.89**
FTO	TG	49.1	48.7	(0.69-1.07)	(0.68-1.06)	**(0.73-1.09)**
	GG	20.3	16.8	0.18	0.15	**0.25**
	G	**43.9**	**41.1**			
	rs1801282					
	CC	78.2	79.0	0.89	0.95	**0.97**
PPAR-γ	CG	20.4	19.3	(0.62-1.23)	(0.65-1.39)	**(0.69-1.361**
	GG**	1.4	1.7	0.53	0.80	**0.87**
	G	**11.8**	**11.3**			
	rs3856806					
	CC	75.8	77.8	0.89	0.89	**0.97**
	CT	23.2	19.7	(0.62-1.23)	(0.62-1.29)	**(0.70 - 1.34)**
	TT**	1.1	2.5	0.53	0.54	**0.86**
	T	**12.7**	**12.3**			
	rs10865710					
	CC	57.8	57.3			
	CG	37.6	36.4	1.02	0.89	**1.06**
	GG**	4.6	6.3	(0.75-1.39)	(0.62-1.29)	**(0.83-1.36)**
	G	**24.5**	**23.4**	0.90	0.54	**0.64**

	SNP rs number Genotype	Controls N=570	FGR cases N=243	Genotype unadjusted OR (95%CI) P value	Genotype adjusted OR* (95%CI) P value	Unadjusted allelic OR (95% CI) P value
CDKN2B-AS1	rs10811661					
	TT	70.8	64.6	1.33	1.26	**1.27**
	TC	25.4	30.8	(0.97-1.84)	(0.91-1.75)	**(0.97-1.67)**
	CC	3.7	4.6	0.08	0.17	**0.09**
	C	**16.5**	**20.0**			
	rs1333049					
	GG	26.1	30.1	0.87 (0.71-1.08) 0.22	0.86 (0.69-1.07) 0.17	**0.87 (0.70-1.08) 0.21**
	GC	48.6	47.3			
	CC	25.4	22.6			
	C	**49.7**	**46.3**			
KCNJ11	rs5215					
	TT	42.9	42.3	1.04	1.07	**1.04**
	TC	44.3	43.6	(0.84-1.30)	(0.85-1.33)	**(0.84-1.30)**
	CC	12.8	14.1	0.72	0.58	**0.71**
	C	**34.9**	**35.9**			
	rs5219					
	CC	42.9	41.8	1.04 (0.83-1.29) 0.40	1.06 (0.85-1.32) 0.61	**1.04 (0.83- 130) 0.73**
	CT	44.1	44.4			
	TT	13.0	13.8			
	T	**35.1**	**36.0**			

Table 2-19: Haplotype frequency of type 2 diabetes genes in FGR cases and control

Haplotype with frequency <5% were pooled as rare haplotypes, P value was using likelihood ratio Chi-square for common haplotypes only. UNPHASED V3.1.5 was used for this analysis.

TCF7L2				
rs7903146 (C/T)	rs7924080 (T/C)	rs12255372 (G/T)	**Cases**	**Controls**
C	T	G	0.50	0.52
C	C	G	0.21	0.17
T	C	T	0.25	0.24
Rare haplotypes			0.04	0.07
P value			**0.35**	
PPAR-γ				
rs1801282 (C/G)	rs3856806 (C/T)	rs10856710(C/G)	**Cases**	**Controls**
C	C	C	0.73	0.73
C	C	G	0.11	0.11
G	T	G	0.09	0.08
Rare haplotypes			0.07	0.08
P value			**0.99**	
CDKN2B-AS1				
rs10811661 (T/C)	rs1333049 (G/C)	-	**Cases**	**Controls**
T	G	-	0.39	0.42
T	C	-	0.41	0.41
C	G	-	0.15	0.08
C	C	-	0.05	0.08
	P value		**0.005**	
KCNJ11				
rs5215 (T/C)	Rs5219 (C/T)	-	**Cases**	**Controls**
T	C	-	0.64	0.65
C	T	-	0.36	0.35
Rare haplotypes			0.02	0.09
	P value		**0.70**	

2.5.6 The effect of T2D variants on maternal BMI and fetal birth weight

The G allele of the FTO SNP rs17817449 was associated with obesity defined by BMI>30Kg/m^2 (OR=1.43 95% CI (1.15 - 1.78), P=0.001) in the control group. Linear regression analysis showed that the G allele of the rs17817449 was significantly associated with maternal BMI regarded as a continuous variable in the control group (Beta=1.15, 95% CI 0.46 - 1.87, P=0.001). The diabetes risk alleles of the markers rs5215 and 5219 in the KCNJ11 gene were associated with lower BMI (P= 0.05 and 0.03 respectively). None of the other loci were associated with BMI. None of the loci tested were associated with baby's birth weight in the control group (Table 2-20).

Table 2-20: Effect of genetic variants on maternal body mass index and baby's birth weight in control groups

Data represent mean (SD) of the trait investigated. The regression coefficient beta, (95% CI), and P values from linear regression are shown. * Adjustment for age, parity, and smoking habits does not alter the findings for maternal BMI, and adjustment for smoking habits did not alter the findings for birth weight (data not shown).

Gene	SNP rs number Genotype	Maternal BMI* (kg/m^2)	Fetal birth weight* (g)
TCF7L2	rs7903146		
	CC	29.7 (6.9)	3455 (618)
	CT	30.0 (7.1)	3467 (562)
	TT	30.0 (6.1)	3473 (637)
	Beta (95%CI)	**0.20 (- 0.59 to 0.1)**	**10.07 (- 65.04 to 85.18)**
	P value	**0.61**	**0.79**
	rs7924080		
	TT	29.9 (6.9)	3467 (590)
	TC	29.8 (6.8)	3454 (619)
	CC	30.2 (7.0)	3476 (572)
	Beta (95%CI)	**0.23 (- 0.47to 0.93)**	**2.89 (- 66.44 to 72.18)**
	P value	**0.52**	**0.94**
	rs12255372		
	GG	29.8 (7.0)	3435 (627)
	GT	30.0 (7.0)	3498 (541)
	TT	30.2 (6.2)	3443 (653)
	Beta (95%CI)	**0.22 (- 0.56 to 1.0)**	**25.23 (- 49.02 to 99.62)**
	P value	**0.57**	**0.51**
FTO	rs17817449		
	TT	28.8 (7.0)	3445 (614)
	TG	30.1 (6.7)	3447 (588)
	GG	31.1 (7.0)	3499 (592)
	Beta (95%CI)	**1.15 (0.45 to 1.86)**	**24.48 (- 45.76 to 94.72)**
	P value	**0.001**	**0.86**
PPAR-γ	rs1801282		
	CC	29.8 (6.7)	3453 (593)
	CG	30.3 (7.4)	3484 (623)
	GG	29.3 (8.3)	3513 (362)
	Beta (95%CI)	**0.27 (- 0.86 to 1.39)**	**30.07 (- 78.57 to 138.71)**
	P value	**0.65**	**0.58**
	rs3856806		
	CC	30.0 (6.8)	3457 (597)
	CT	29.7 (7.1)	3495 (595)
	TT	29.0 (8.4)	3112 (637)
	Beta (95%CI)	**- 0.32 (- 1.46 to 0.82)**	**1.04 (- 106.76 to 108.84)**
	P value	**0.58**	**0.98**
	rs10865710		
	CC	30.0 (6.7)	3455 (585)
	CG	30.0 (7.2)	3479 (625)
	GG	27.4(5.4)	3442 (449)
	Beta (95%CI)	**- 0.62 (- 1.48 to 0.25)**	**4.47 (- 72.83 to 95.78**
	P value	**0.16**	**0.78**

Gene	SNP rs number Genotype	Maternal BMI* (kg/m^2)	Fetal birth weight* (g)
CDKN2B-AS1	rs10811661		
	TT	29.9 (7.0)	3463 (595)
	TC	29.7 (6.6)	3434 (625)
	CC	30.9 (6.8)	3569 (438)
	Beta (95%CI)	**0.09 (- 0.19 to 0.20)**	**- 1.91 (- 89.33 to 85.52)**
	P value	**0.85**	**0.96**
	RS1333049		
	GG	30.2 (7.2)	3418 (602)
	GC	29.8 (6.8)	3483 (578)
	CC	29.9 (6.8)	3463 (626)
	Beta (95%CI)	**0.10(- 0.59 to - 0.805)**	**22.66 (- 46.22 to 91.55)**
	P value	**0.77**	**0.52**
KCNJ11	rs5215		
	TT	30.5 (7.0)	3477 (562)
	TC	29.6 (7.0)	3424 (629)
	CC	28.8 (6.2)	3519 (588)
	Beta (95%CI)	**- 0.75 (- 1.5 to - 0.01)**	**0.12 (- 72.22 to 72.45)**
	P value	**0.05**	**0.99**
	Rs5219		
	CC	30.5 (7.0)	3481 (563)
	CT	29.6 (7.0)	3427 (636)
	TT	28.9 (6.2)	3520 (584)
	Beta (95%CI)	**- 0.86 (- 1.62 to - 0.09)**	**- 1.29 (- 73.92 to 71.35)**
	P value	**0.03**	**0.98**

2.6 Discussion

2.6.1 Type 2 diabetes genes in preeclampsia case control study

It is widely reported that diabetic women are at increased risk of preeclampsia (Rosenberg et al., 2005), and also that women with a history of preeclampsia are at a higher risk for developing diabetes (Lykke et al., 2009, Libby, 2007) and coronary artery disease (Ray et al., 2005, Lykke et al., 2009) later in life. This led to the term "metabolic syndrome of pregnancy" being applied to preeclampsia. A large body of evidence is currently available indicating a genetic basis for preeclampsia, diabetes and CAD.

Variants in TCF7L2, FTO, PPAR-γ, CDKN2B-AS1, and KCNJ11 genes were reproducibly associated with T2D in candidate gene studies and GWAS in Caucasian and other ethnic groups, marking these genes as authentic susceptibility genes for T2D. While variants of FTO and PPAR-γ genes confer risk for T2D by altering insulin sensitivity in the peripheral tissues, mainly adipose tissue (Honka et al., 2009, Do et al., 2008), variants of TCF7L2, CDKN2B-AS1 and KCNJ11 are considered to confer risk for T2D via changes in pancreatic beta cell function, mainly insulin secretion (Lyssenko et al., 2007, Schwanstecher et al., 2002). These genetic variants are also associated with metabolic changes which precede overt T2D such as hyperglycaemia and hyperinsulinaemia and markers of metabolic syndrome like dyslipidaemia. Indeed, variants of PPAR-γ, FTO, and CDKN2B-AS1 genes have been associated with CAD, and atherosclerosis (McPherson et al., 2007, Samani et al., 2007, Doney et al., 2009, Ridker et al., 2003).

Using 11 SNPs in 5 confirmed T2D genes (Table 2-7), this study has tested the hypothesis of shared genetic origins between preeclampsia and diabetes in a case control study. The association was tested in a collection of women who were at higher risk for preeclampsia (Table 2-10). By including women with diabetes, hypertension and obesity which are known risk factors for preeclampsia, the study does not exclude genetic factors related to these disorders which might be shared with preeclampsia. This is in contrast with previous investigations of women with no medical or obstetric risk factors, which may select against those with susceptibility genes shared by preeclampsia and related medical conditions.

In the present study, there was no evidence for any association between these T2D genetic variants and preeclampsia (Table 2-12). To the best of our knowledge, this is the largest study, in terms of sample size and number of variants, to assess the role of common variants in genes predisposing for T2D and their putative role in preeclampsia. This is also the first study to investigate variants of TCF7L2, FTO, and KCNJ11 in relation to preeclampsia.

Similar to the findings of this study, a Finnish group found no association between PPAR-γ Pro12Ala (rs1801282) and preeclampsia in a small study of 133 cases and 115 controls (Laasanen et al., 2002). A lack of association between the Ala (G allele) variant of SNP rs1801282 of the PPAR-γ and preeclampsia was also reported in a recent study by a German group (Wiedemann et al., 2010). In both of these studies the power to detect a genetic effect was limited by both the sample size and the low MAF of the variant investigated. Although this study has 85% power to detect an OR of 1.5 at P value <0.05 for this specific variant with MAF 0.12, power to detect OR as low

as 1.2 and 1.1 (the effect size reported for T2D for this specific variant) is limited (50% and 25%, respectively). The results of this study suggest that the Ala allele (G allele of SNP rs1801282), the diabetes protective marker that is reportedly associated with higher insulin sensitivity and low BMI, is a risk allele for preeclampsia (OR 1.15, 95% CI 0.87-1.53) although this was not significant (Table 2-12). Interestingly in the Finnish study the Ala allele was associated with early age at delivery and was more common in severe than milder forms of preeclampsia (Laasanen et al., 2002).

Polymorphisms in the promoter of the PPAR-γ (C-681G; rs10856710) and in exon 6 of the gene (C1431T; rs3856806) have been associated with higher body weight and plasma LDL concentrations respectively (Doney et al., 2002, Meirhaeghe et al., 2005). The effect of the Ala allele on obesity was found to be altered by these 2 variants which are in LD with it. A haplotype formed of the 3 polymorphisms was associated with the metabolic syndrome (Meirhaeghe et al., 2005). However haplotype analysis in the present investigation revealed no association with preeclampsia, or obesity (Tables 2-13 and 2-20).

PPAR-γ is a ligand-activated transcription factor playing an important role in adipocyte differentiation and glucose homeostasis. PPAR-γ also has anti-inflammatory effects through reducing the production of pro-inflammatory cytokines after forming complexes with Nf-KB. PPAR-γ agonists increase the size of LDL particles and decrease the number of atherogenic small dense LDL particles. It is also reported that it has a blood pressure lowering effect probably through an inhibitory effect on thromboxane A2 and the RAS system (Sugawara et al., 2010). PPAR-γ also exhibited anti-inflammatory on human gestational

tissues as both natural (15dPGJ2) and synthetic ligands (troglitazone) were shown to significantly decrease lipopolysaccharide stimulated secretion of proinflammatory cytokines from gestational tissues (Lappas et al., 2002).

In vitro studies have shown that the PPAR-γ Ala12 allele of SNP rs1801282 decreases PPAR-γ transactivation of reporter genes and the -681G allele of SNP rs10856710 is associated with lower PPAR-γ promoter activity (Meirhaeghe et al., 2003) suggesting that carriers of these alleles might have a lower PPAR-γ - mediated activation of target genes than non carriers. So it is biologically feasible that the Ala allele may confer risk for preeclampsia through these functional effects, and larger studies will be needed to exclude a small effect size.

The CDKN2B-AS1 gene encodes long non-coding RNA with no specific function attributed to it. It is suggested that polymorphisms in this region are implicated in many diseases through altering the expression of CDKN2B-AS1 and the nearby genes such as CDKN2A/AB, tumour suppressor genes and cell cycle regulators (Cunnington et al., 2010). CDKN2B-AS1 in the region on chromosome 9p21 harbours 2 SNPs in 2 adjacent LD blocks (Figure 2-6). The first is associated with T2D (rs10811661) and the second (rs1333049) is associated with CAD, with no overlap in the risk conferred by either of the SNPs towards T2D and CAD (Broadbent et al., 2008, Gori et al., 2010). In the current study no association was detected between either of the 2 SNPs and preeclampsia (Table 2-12 and Table 2-13). This region on chromosome 9p13-21 was linked to preeclampsia in a Finnish family study (Laivuori et al., 2003) In a recent case control study, extensive investigation of this region testing 23 SNPs

including the two tested in the present study, showed two markers (rs7044859 and rs496892) to confer risk for preeclampsia in Finnish women (Peterson et al, submitted for publication). A subset of samples used in the present study and an Australian family study did not reproduce these finding. The number of UK samples was small, but an almost 2 fold increase in sample size in the present study failed to confirm the findings from the Finnish women. It is possible that this was a spurious finding, or that the causal variant in this region has a pattern of LD with the markers investigated which differs in UK and Finnish populations, and consequently was not tagged by the 2 markers used in the present study.

Surprisingly, the G allele of marker rs17817449 of the FTO gene, the allele that was reportedly associated with higher BMI, T2D, CAD, and hypertension and their intermediate metabolic measures exhibited a borderline protective effect against preeclampsia (OR 0.84, 95% CI 0.708-1.0, P=0.05), (Table 2-12). This variant (or variants in LD with it) not only increases T2D risk but also increases inflammatory mediators like CRP and is associated with an atherogenic lipid profile (Doney et al., 2009), all of which are related to the pathogenesis of preeclampsia. FTO gene expression was correlated with the expression of hypoxia inducible factor-1a (HIF-1a) in adipose tissue, and with high-mobility group protein B1 (HMGB1), a cytokine mediator of inflammation (Lappalainen et al., 2010) both of these pathways have a pivotal role in preeclampsia. However this apparent protective effect is spurious and confounded by the higher BMI in controls than in preeclampsia cases as evidenced by elimination of this effect after adjusting for BMI (OR 0.89, 95% CI 0.74-1.09, P=0.27), (Table 2-12). In agreement with previous studies the G allele was associated with

obesity, defined by BMI>30Kg/m^2 (OR=1.15, 95% CI 0.46 - 1.87, P=0.001) and BMI (beta =1.43; 95% CI 1.15-1.78, P=0.001), (Table 2-20).

No association could be detected between preeclampsia and variants in the TCF7L2 or KCNJ11 genes (Table 2-12). Risk conferred by these 2 genes for T2D is primarily related to beta cell function, reduced insulin secretion and processing, while insulin resistance, hyperglycaemia and hyperinsulinaemia are the pathological feature of preeclampsia.

The lack of association between T2D genetic variants and preeclampsia may be due to inadequate power to detect small genetic effect sizes below 10%, and replication is required in larger studies to narrow down the 95% CI. No attempt was made to capture all variation in each gene, as the selected SNPs were those shown to alter T2D risk. The possibility remains that other genetic variants carried on these genes can affect the risk of preeclampsia. Each of these genes has pleiotropic functions that might be tissue or organ specific so that variants that confer risk for T2D may have a different effect on the risk of preeclampsia. It is also possible that preeclampsia and diabetes do not share the same genetic risk factors, and that the bidirectional risk between them is mainly due to shared biochemical and pathological changes. It is worth noting that the study pooled 3 cohorts, VIP, NottFGR and NottPE. In the later 2 collections, both cases and controls were recruited from otherwise healthy women with no known risk factor for preeclampsia or FGR. By contrast, in the VIP study recruitment was from women with known medical and/or obstetric risk factors for preeclampsia. In addition to this difference between the cohorts due to risk factor enrichment in the VIP study, maternal age and BMI were higher while frequency of nulliparity

was lower in VIP study. So by pooling these 3 collections of women elements of phenotypic heterogeneity have been introduced, which may reflect underlying genetic heterogeneity. This would weaken the power of the study if genetic and pathophysiological determinants of preeclampsia differ between women with and without high risk factors. However the detection of markers of endothelial dysfunction early in pregnancy in women who subsequently develop pre-eclampsia casts doubt on the two stage hypothesis of pre-eclampsia (defective placentation followed by endothelial dysfunction). Moreover, this study was based on the hypothesis that preeclampsia and the cardio-metabolic diseases are parts of a continuum of disorders with possible shared genetic risk factors for both conditions. Pooling of the 3 cohorts is therefore justifiable, and meets the aim of the present study to discover shared genetic risk factors between T2D and preeclampsia. Indeed, attempts to analyze samples from VIP collection separately yielded the same findings. No attempt was made to analyze the other 2 groups as even if pooled together the sample size is insufficient for an adequately powered study.

In conclusion, none of the 11 loci on TCF7L2, FTO, PPAR-γ, KCNJ11, or CDKN2B-AS1 genes showed variations in the frequency between preeclampsia cases and controls, and therefore the hypothesis that the bidirectional risk between T2D and preeclampsia is of genetic origin was not confirmed.

2.6.2 Type 2 diabetes genes in FGR case control study

Many studies have investigated the effect of fetal genotype at T2D loci on birth weight in an attempt to provide support for the fetal insulin hypothesis, advanced as an explanation for the interaction between low birth weight and T2D in adult life. This hypothesis states that the genetic variants that lead to T2D in adult life due to reduced insulin secretion and/or increased insulin resistance will lead to low birth weight (Frayling and Hattersley, 2001) as insulin is a major regulator of fetal growth, and fetal insulin levels are positively correlated with birth weight (Verhaeghe et al., 1993). Previous studies have indicated that mothers who give birth to small sized babies developed insulin resistance and / or T2D later in their lives (Hypponen et al., 2003, Lawlor et al., 2002). These findings imply that those mothers, potential carriers of T2D diabetes risk alleles, were at risk of having growth restricted pregnancy. As 50% of fetal genes are maternally inherited, we investigated the effect of maternal genotype at T2D susceptibility loci on the risk of having a growth restricted pregnancy.

There was a marginal association between the C (diabetes protective) allele of the rs10811661 of the CDKN2B-AS1 gene and the risk for FGR (OR = 1.33, 95% CI 0.97-1.84, P =0 .08) that was abolished after adjusting for the risk factors and smoking habit (OR = 1.26, 95% CI 0.91-1.75, P=0.17), (Table 2-18). Haplotype analysis using the 2 loci in this region showed that the haplotype (CG) formed of the diabetes protective allele and CAD protective allele was more common in FGR cases than controls (14.7% vs8.6%, P=0.005), (Table 2-19). Mothers who carried the CG haplotype were at double the risk of having a growth restricted pregnancy (OR 1.98, 95% CI 1.33- 2.95, P = 0.0008) compared to mothers who carried the TG haplotype (the reference haplotype formed of the

two major alleles). As the variant rs1333049 alone showed no alteration in the risk of FGR it seems that interaction between the two loci (C allele of the rs10811661 and the G allele of rs1333049) lies behind this association.

Jarinova et al (2009) showed that the C allele of the SNP rs1333045 (in complete LD with rs1333049) had an enhancer effect on the expression of luciferase in a reporter gene construct and was associated with a 2 fold increase in the expression of the CDKN2B-AS1 in primary aortic smooth muscle cells (Jarinova et al., 2009). In this study, CDKN2B-AS1 expression was correlated with CDKN2B gene expression. It has been suggested that variants on chromosome 9p21 affect the expression of CDKN2B-AS1 and CDKN2A/2B genes (Burd et al., 2010, Cunnington et al., 2010). Reduced expressions of CDKN2A/2B, which are tumour suppressor genes and cell cycle regulator genes, have been implicated in atherosclerosis and apoptosis (Jarinova et al., 2009). This is the first study to investigate the CDKN2B-AS1 T2D variants and the risk of FGR. Animal studies indicated that deletion of the CDKN2B-AS1 locus in mice was associated with higher birth weight in offspring (Visel et al., 2010); these studies suggest a putative role for CDKN2B-AS1 in growth regulation.

Interestingly, neither of the 2 variants in the CDKN2B-AS1 gene was associated with birth weight in the control group (Table 2-20). These findings are in line with many previous studies (Freathy et al., 2009, Freathy et al., 2010, Andersson et al., 2010) and contradict a single study that indicated high birth weight in relation to fetal genotype of the diabetes risk allele of rs10811661 (Pulizzi et al., 2009). The authors did not report on the maternal genotype or glycaemic status.

The suggestion that this locus of the CDKN2B-AS1 gene is a potential candidate for FGR but is not related to birth weight in healthy pregnancies may reflect differences in the genetic factors governing variations in birth weight and those determining growth restriction. FGR should not simply be regarded as the lowest extreme of birth weight; it is the result of aberrant pathophysiological mechanisms in pregnancy, and possibly an altered genetic architecture.. Many healthy babies are small at birth simply because they are constitutionally small; others show features of growth restriction even though their birth weight lies within the normal range. It is for this reason that defining FGR pregnancy is challenging. Haplotype analysis with birth weight as a quantitative trait may be considered although it carries some statistical limitations. The difficulty is that haplotypes are not observed; instead, they are inferred. Case control studies rely mainly on estimated haplotype proportions among cases and controls, without clear haplotype assignment for individuals (Balding, 2006). When analysing quantitative traits it is necessary to assign haplotypes to individuals, which cannot be done unequivocally in those who are heterozygous at more than one SNP. For this reason, single locus analysis is more powerful than haplotype analysis for a quantitative trait such as birth weight, even when the locus is not causative but is in LD with the causal locus (Schaid, 2004).

There was no evidence for an association between the maternal T2D risk alleles in the TCF7L2 gene and FGR (Table 2-18 and Table 2-19) or birth weight (Table 2-20) in healthy pregnancies. This study is the first to investigate maternal T2D risk alleles in FGR. Previous studies reported no association between fetal genotype at TCF7L2 rs7903146 and smallness for gestational age (Cauchi et al., 2007b, Mook-Kanamori et al., 2009). Similarly, the study by Cauchi et al

(2007b) reported no effect of maternal or fetal genotype at the TCF7L2 rs7903146 locus on birth weight (Cauchi et al., 2007b). In contrast, a large study by Freathy et al, (2007) found that each maternal T2D T allele at this locus was associated with a 40g increase in birth weight (Freathy et al., 2007).

There was also no association between maternal genotype at FTO, PPAR-γ, and KCNJ11 variants and FGR, or birth weight in the control group (Table 2-18 to 2-20). Previous well powered studies of fetal genotype have not detected any association between T2D risk alleles of PPAR-γ, KCNJ11, or FTO genes (Bennett et al., 2008, Mei et al., 2010, Kilpelainen et al., 2011).

The reported lack of association between with many T2D variants and FGR can be related to many factors. First is the sample size. The sample size used in this study had limited power to detect small effects in the range of 10-20% (Table 2-3). A sample size of 1000 cases and 2000 controls is required to detect a 20% change in the risk of FGR at 80% power, P =0.05, for variants with MAF of 0.2. Second is that maternal genotypes at T2D risk alleles might be expected to have similar effects on the risk of fetal growth restriction as fetal genotypes, but only if it were possible to test them independently of their metabolic effects on the mother. The impact of these loci on maternal glycaemic control makes it difficult to separate the influence of genes transferred from mother to offspring from that of the maternal environment, and could result in apparently opposing effects of maternal and fetal genotype. Consequently, it cannot be excluded that, when the risk allele is present in both the mother and the baby, small effects could be masked. Previous studies have shown that maternal genotype modulated the effect of fetal genotype on birth weight. An important example is variants of the

GSK gene (Weedon et al., 2005, Freathy et al., 2007) and TCF7L2 (Freathy et al., 2007). Lastly we pooled 3 collections of women , the NottFGR cases were very tightly defined; NottmPE FGR cases almost certainly had growth restricted pregnancies, on the basis that PE and FGR share similar placental mechanisms; VIP FGR cases, unless they had PE during the VIP pregnancy, are SGA, and only a minority of these will have FGR. This might create phenotypic heterogeneity that reduces the statistical power of the study and hampers the effort to detect and association.

In summary, this study suggested a region on chromosome 9p21, at the CDKN2B-AS1 gene locus, as a potential risk locus for FGR, which merits further investigation. There was no evidence for any association of the T2D risk alleles of TCF7L2, FTO, PPAR-γ, or KCNJ11 with fetal growth restriction.

3 Role of epidermal growth factor genetic variants in preeclampsia and fetal growth restriction

3.1 Introduction

Epidermal growth factor (EGF) is an autocrine/paracrine modulator of cell growth, differentiation, and function in many tissues of mammalian species (Fisher and Lakshmanan, 1990). It was first discovered in the mouse submaxillary gland, and later in human, where it was initially named urogastrone (Bell et al., 1986).

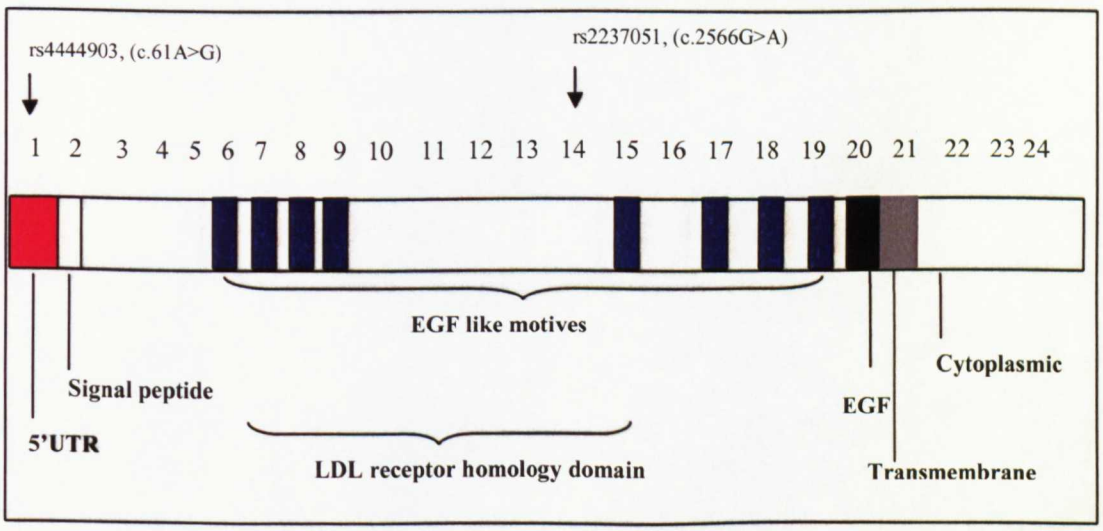
EGF is secreted as a membrane bound protein (prepro-EGF) formed of 1207 AA. This precursor protein is subjected to enzymatic cleavage of 970 and 180 AA from the amino-terminal and carboxy-terminal ends respectively, to give the mature EGF. Eight cysteine rich repeats were identified in this precursor protein and termed EGF-like motifs. The prepro-form of EGF shows similarity with LDL receptors, and it was suggested that it may act as a receptor for an unidentified ligand especially in the kidney. PreproEGF was found to be biologically active, binding to and activating the EGF receptor (EGFR) and maintaining cellular growth in culture (Mroczkowski and Reich, 1993).

The mature EGF protein has a molecular weight of 6.3 kDa and is formed of 53 AA with 3 disulfide bonds between 6 cysteine residues. These disulfide bonds are conserved along with 66% of the AA sequence between human and mouse EGF. EGF is a member of a growth factor family that shares the EGF-like motif and includes transforming growth factor-α (TGF-α), betacellulin, amphiregulin, heparin binding EGF, and epiregulin.

The EGF gene occupies around 100Kb of chromosome 4 between bases 111053449- 111152868, forming one large LD block. Human EGF has only one known transcript and is expressed from a TATA box promoter located 42 bp upstream of the transcription start site. The EGF gene is formed of 24 exons. The 1st and one of the largest exons of the EGF gene contains the 442 bp 5' untranslated region (5'UTR) and the start of the coding sequence which encodes the signal peptide responsible for targeting the protein to the cell membrane. The mature EGF is encoded by exons 20 and 21. Exons 6-9 and 15-19 code for the 8 cysteine rich EGF-like motifs, with sequence similarity to LDL receptors (Figure 3-1).

Figure 3-1: Genomic structure of the epidermal growth factor gene

The EGF gene is formed of 24 exons (numbered). The encoded protein domains are also shown. The 2 SNPs included this study are arrowed.



A large body of biological and genetic evidence suggests EGF as a candidate gene for preeclampsia and FGR. This chapter includes 4 sections: the first focuses on the role of variants of the EGF gene as risk factors for preeclampsia

and FGR; the second and third sections explore the functional effects of the SNPs rs2237051 and rs4444903 on EGF gene expression. The last section of this chapter describes how these variants alter the expression of EGF and EGFR in the placenta of normal pregnancy and pregnancy complicated by preeclampsia and FGR.

3.2 Aims of the study

- To investigate the risk conferred by 2 EGF genetic variants in preeclampsia and FGR in a case-control study.
- To explore the mechanisms whereby these variants affect EGF expression.
- To do a preliminary investigation how these variants affect the expression of EGF and EGFR in the placenta of normal pregnancy and pregnancy complicated by preeclampsia and FGR.

3.3 Genetic association study

3.3.1 Introduction

Considering the close proximity of the EGF gene (4q25-27) to the loci on chromosome 4 linked to preeclampsia (4q32) (Harrison et al., 1997, Laivuori et al., 2003) and birth weight (4q31) (Arya et al., 2006) and the biological properties of EGF in trophoblast function, EGF is considered as a potential candidate gene for preeclampsia and FGR.

3.3.1.1 Epidermal growth factor and placental development

EGF is a growth factor that is expressed in a wide variety of tissues and organs including the reproductive tissues (Hofmann et al., 1991). Multiple studies have shown that EGF is actually expressed in placental tissues. By immunostaining, EGF expression was noted in decidua and placenta throughout gestation. In decidua it was more prominent around the spiral arteries; in early placenta it was abundant and localized to syncytiotrophoblast, while light to moderate staining was localized to both trophoblast layers in late pregnancy (Hofmann et al., 1991). EGF mRNA and protein were also detected in primary trophoblast cells along with its receptor (Amemiya et al., 1994). Moreover EGF was detected in culture media of early placental explants (Maruo et al., 1995). EGF in the placenta was suggested to be of maternal and fetal origin. Bass et al (1994) failed to show expression of EGF from either 1st trimester and term cytotrophoblast or villous fibroblast in culture media using radioimmunoassay, although addition of EGF to culture media increased cytotrohoblast invasiveness. This study suggested that it is EGF, mainly of maternal origin, which regulates trophoblast properties during early pregnancy, when the embryo also expresses

EGF receptors (Bass et al., 1994). This lack of expression is probably related to the short duration of culture (2 days) that may have resulted in EGF production being below the detection level of the method applied (15 pg/7 million cells).

EGF has been shown to increase 1st trimester CT trophoblast cell proliferation in culture (Li and Zhuang, 1997, Maruo et al., 1995). EGF also induces trophoblast invasion. Addition of EGF caused dispersion and formation of pseudopodia in 1st and 2nd trimester CT trophoblast cells (Bass et al., 1994, Staun-Ram et al., 2004). The enhancement of invasive properties was also evident in a model of extravillous trophoblast cells SGH-PL4 (LaMarca et al., 2008). EGF appears to exert two different actions independently in a stepwise manner on the proliferation and functional differentiation of trophoblast. EGF enhances CT trophoblast proliferation in culture of very early (4-5 week) placental tissues, without affecting the ability to secrete hCG and hPL. This is in contrast to the effect of EGF observed on older placental tissues in culture (6-12 weeks), in which EGF stimulated hCG and hPL secretion without affecting proliferation (Maruo et al., 1992). Morrish et al (1997) demonstrated that EGF enhanced syncytium formation from term CT trophoblast in culture (Morrish et al., 1997). Activation of MAPK11/14 (P38) by EGF was implicated in the differentiation of CT trophoblast into ST trophoblast (Johnstone et al., 2005).

Control of trophoblast invasiveness by EGF involves activation of multiple cellular pathways. EGF stimulates MMP-2 and -9 secretion from trophoblast cell lines in a differential, dynamic manner in which MMP-2 is dominant in early 1st trimester and both MMP-2 and 9 late in the 1st trimester (Staun-Ram et al.,

2004, Qiu et al., 2004). It also induces urokinase-type plasminogen activator and plasminogen activator inhibitor-1 activity in trophoblastic cells, thereby inducing cell invasion. Besides altering the extracellular matrix degrading enzymes, EGF also changes the extracellular matrix composition. EGF was found to alter trophoblast migration/invasion through induction of α2 integrin expression (Nakatsuji et al., 2003) and down regulation of connexin 40 (Wright et al., 2006).

Apoptosis, an important process in the maintenance of the ST trophoblast layer during healthy pregnancy, is shown to be exaggerated in placenta from preeclampsia and FGR complicated pregnancy (Ishihara et al., 2002, Crocker et al., 2003). EGF was found to reduce the apoptotic process in a trophoblast cell lineage exposed to hypoxia, reactive oxygen species or cytokines (Humphrey et al., 2008, Moll et al., 2007, Smith et al., 2002).

3.3.1.2 EGF and fetal growth

EGF is regarded as an important factor not only for decidualization, implantation and placental function during the course of pregnancy but also for fetal growth, organ development and maturation.

In animal models, the body weight of fetuses whose mothers had undergone sialoadenectomy (removal of the submaxilary gland, a major source of EGF) or had been administered EGF antibody were significantly lower than that of controls (Odaka et al., 1993, Masuyama et al., 1996). The offspring of rats administered EGF during pregnancy were larger than controls (Masuyama et al., 1996). The fetomaternal amino acid concentration ratio (fetal blood/maternal blood) was higher in the EGF-treated group than in the control group

(Masuyama et al., 1996). EGF promoted amino acid transport in placental microvillus explants and pregnant rats, influencing fetal growth as a result (Bonkobara et al., 2002). Moreover, maternal EGF deficiency during the latter half of pregnancy led to reduced placental expression of the glucose transporter GLUT3 and selectively impaired the transplacental glucose supply, resulting in intrauterine growth retardation in mice. These findings suggest that EGF may play an important role in regulating placental function and fetal growth (Kamei et al., 1999).

EGF has also been implicated in growth and maturation of different fetal tissues and organs. Intestinal epithelium exhibits its highest EGF binding capabilities between 8 and 12 weeks of fetal life, which is a critical period of intestinal morphogenesis and differentiation. EGF exerts important effects on epithelial cell proliferation of human jejunum and duodenum during the fetal period (Menard et al., 1988, Beaulieu et al., 1985). In support of these findings, supplementation of amniotic fluid by EGF normalizes fetal weight and intestinal proliferation in growth restricted fetal rabbits (Buchmiller et al., 1993, Cellini et al., 2004).

Epidermal growth factor also influences fetal lung development. It was demonstrated that the weight of the lungs was low in the offspring of EGF-immunized rats. A picture suggesting a delayed lung maturation resulting in mild respiratory distress syndrome was also noted (Raaberg et al., 1995). This was in agreement with other studies which reported the positive effect of EGF on enhancing maturation of fetal lung in animal models, by increasing the production of various surfactant components (Ma et al., 2009).

By stimulating the prostaglandin H2 synthetase enzyme in human amnion cells (Casey et al., 1988) and up regulation of cyclooxygenase 1 (COX1) and COX2 in amniotic fluid and umbilical cord, EGF was suggested to maintain a steady state of blood supply to the growing fetus by maintaining the tone of the umbilical blood vessels (Rao et al., 1995). In addition, EGF may constrict umbilical blood vessels immediately after delivery to prevent bleeding of the fetus.

3.3.1.3 Epidermal growth factor levels during pregnancy

Many studies have measured and traced EGF levels in maternal urine and in amniotic fluid throughout pregnancy. Maternal EGF urinary levels increased from early pregnancy until mid term, when they declined to a level at term similar to non-pregnant controls (Moharam et al., 1992, Hofmann and Abramowicz, 1990, Watanabe, 1990). By contrast amniotic fluid and cord serum EGF concentrations were higher near term than mid-gestation (Hofmann and Abramowicz, 1990, Ichiba et al., 1992). The lack of correlation between amniotic fluid and maternal urinary EGF concentrations suggests that probably there is a different source of EGF in the two compartments. In contrast, other studies have demonstrated that EGF levels in both amniotic fluid and urine increased with advancing gestation, and reported that levels in FGR cases were lower compared with normal pregnancy (Watanabe, 1990, Lindqvist et al., 1999).

3.3.1.4 Polymorphisms of the EGF gene and the risk of preeclampsia and FGR

Two SNPs in the EGF gene have attracted interest. The first one, located in the 5'UTR of the gene (c.61A>G or rs4444903), has a MAF (G allele) of 0.44 in Caucasians and is a potentially functional SNP. Shabahzi et al (2002) demonstrated that the G allele was associated with an increased risk of malignant melanoma (OR 2.7, 95% CI 2.9-4.0, P=0.0001) and showed more EGF production from peripheral blood mononuclear cells with the G allele compared to the A allele (Shahbazi et al., 2002). This SNP has also been associated with other types of malignancies. For example the G allele conferred higher risks for gliomas (OR, 1.32; 95% CI, 1.04-1.67), and glioblastomas (OR, 1.47; 95% CI, 1.02-2.10) (Costa et al., 2007). It also was associated with an increased risk for hepatocellular carcinoma, (Tanabe et al., 2008), and oesophageal carcinoma (Lanuti et al., 2008), but was not associated with breast cancer (Wang et al., 2008).

The second SNP in the EGF gene is located in exon 14, with a MAF of 0.37 in Caucasians. It is a non-synonymous SNP that leads to a change in the encoded amino acid from methionine to isoleucine (c.2556G>A or rs2237051) (Semina et al., 1996). This substitution does not affect the mature EGF encoded by exons 20 and 21, but it can potentially affect the properties and functionality of the preproform of the EGF. However, as they are both neutral, non-polar AA, methionine and isoleucine share similar physiochemical properties, lessening the probability that this variation alters the properties of the protein.

The impact of these EGF variants on birth weight variability has been previously investigated (Dissanayake et al., 2007). They showed that the maternal G-A

haplotype formed of the minor alleles of markers rs4444903 and rs2237051 was associated with low birth weight in healthy pregnancies in Caucasian women and was transmitted more frequently from parents to babies affected by FGR. They subsequently reported that the GA and GG genotypes of the marker rs2237051 were associated with higher risk for preeclampsia in Sinhalese women compared to the AA genotype (OR 1.7 (95% CI 1.0-2.8) and OR 1.8, (95% CI 1.0-3.2) respectively (Dissanayake et al., 2009).

Based on the aforementioned biological and genetic findings, the aims of this study were to investigate the maternal genetic variants of EGF as risk factors for preeclampsia and FGR in a case control study, and to test the effect of these variants on the maximum blood pressure measured during pregnancy in healthy pregnant women, and birth weight variability of their offspring.

3.3.2 Methods

Patients and samples used for this investigation were identical to those described in the previous chapter (Section 2.4.1). The subset of 105 cases and 90 controls from the Nottingham FGR collection were also included in the study carried out by Dissanayake et al (Dissanayake et al., 2007).

A total of 673 women with normal pregnancies, 448 women with preeclampsia and 234 mothers with FGR pregnancies were genotyped at the EGF SNPs c.61A>G (rs4444903) and c.2556 G>A (rs2237051) by Kbiosciences using the KASPr technique.

3.3.3 Results

Demographic and clinical characteristics of the study population were presented in chapter 2 (Table 2-8, Table 2-9, Table 2-14 and Table 2-15).

The call rate for the genotyping of markers rs4444903 and rs2237051 was 98.2% and 98.6% respectively. Both were in HWE within individual plates, the control group and within the whole sample set (P >0.05). The MAF was 0.41 and 0.37 for rs4444903 and rs2237051 respectively, similar to those reported in the HapMap project for the European population (Table 3-1). In the study by Dissanayake et al., (2007) SNP c.61A>G (rs4444903) was genotyped by restriction fragment length polymorphism and c.2556 G>A (rs2237051) by MS-PCR. 100% concordance was observed between the results for the 200 samples genotyped by both Dissanayake et al. and the current study.

The 2 markers were in linkage disequilibrium (r^2 = 0.81). The A-G and G-A haplotypes formed of the 2 major and 2 minor alleles respectively accounted for 95% of the observed haplotypes (Table 3-4).

Table 3-1: Genotyping quality control indices

Chr = Chromosomal location based on HapMap release#24

SNP	Location	Function	Major/ Minor allele	MAF	Success rate %	Hardy Weinberg (P value)
rs4444903	Chr.4 111.053.559	5'UTR	A/G	0.42	98.2	0.41
rs2237051	Chr.4 111.120.647	Exon 14 preproEGF	G/A	0.39	98.6	0.32

3.3.3.1 Association of EGF variants with preeclampsia

The G allele of rs4444903 was less frequent in preeclampsia cases (0.38) compared to controls (0.42). The allelic OR was 0.84, (95% CI 0.71-1.00, P= 0.06). Comparison of the genotype at this locus showed a protective effect for the GG genotype of marginal significance (OR 0.84, 95% CI 0.70- 1.00, P= 0.05). This protective effect was not statistically significant after adjusting for parity, smoking habit, BMI and the risk factors including chronic hypertension, diabetic status, chronic renal disease, abnormal uterine artery Doppler, and multiple pregnancies (OR 0.88, 95% CI 0.73-1.07, P = 0.20), (Table 3-2)

There was no significant association between the A allele (OR 0.90, 95% CI 0.76-1.07) and AA genotype (OR 0.90, 95% CI 0.76-1.07) of rs2237051, before or after adjustment for confounding factors (Table 3-2).

The haplotype G-A formed of the minor alleles of the 2 variants was represented more frequently in controls (38%) than preeclamptic cases (35%), but this was not statistically significant (P = 0.11; Table 3-4).

3.3.3.2 Association of EGF variants with FGR

In the FGR case-control comparison, there was no significant association between the G allele of rs4444903 (OR 0.84, 95% CI 0.68- 1.05, P = 0.13). The allelic OR for the A allele of rs2237051 was also non-significant (OR 0.92, 95% CI 0.74-1.145, P = 0.45). The same was observed when genotype was analyzed both before and after adjustment for smoking habit, and the risk factors including chronic hypertension, diabetic status, chronic renal disease, and abnormal uterine

artery Doppler (Table 3-3). We did not adjust for parity, maternal age, gestational age or maternal BMI in this analysis, as those are corrected in the calculation of the GROW centile used in the identification of cases and controls. Multiple pregnancy was excluded by definition.

The haplotype G-A formed of the minor alleles of the 2 variants was represented more frequently in controls (38%) than FGR cases (35%), but this was statistically non significant (P = 0.24; Table 3-4).

Table 3-2: Genotype and allele distributions for EGF variants and corresponding odds ratios for preeclampsia

Genotype frequency was calculated using cross tabs, and genotype OR was calculated using binary logistic regression, both using SPSS v.16. * Adjusted for parity, smoking habit, BMI and the risk factors including chronic hypertension, diabetic status, chronic renal disease, abnormal uterine artery Doppler, and multiple pregnancies. Allelic OR was calculated using UNPHASED software release 3.1.5.

Locus	Genotype			Allele	Genotype unadjusted OR (95%CI) P value	Genotype adjusted OR* (95%CI) P value	Unadjusted allelic OR (95% CI) P value
rs4444903	**AA**	**AG**	**GG**	**G**			
Preeclampsia (N=448)	38.0	49.1	13.0	**0.38**	**0.84 (0.70- 1.00) 0.05**	0.88 (0.73-1.07) 0.203	**0.84 (0.71-1.00) 0.06**
Healthy pregnancy (N=673)	33.3	50.2	16.5	**0.42**			
rs2237051	**GG**	**GA**	**AA**	**A**			
Preeclampsia (N=448)	139.5	47.0	13.4	**0.37**	0.90 (0.76-1.07) 0.25	0.94 (0.78-1.14) 0.54	0.90 (0.76-1.07) 0.24
Healthy pregnancy (N=673)	37.0	47.3	15.8	**0.39**			

Table 3-3: Genotype and allele distributions of the EGF variants and corresponding odds ratios for fetal growth restriction (FGR)

Genotype frequency was calculated using cross tabs, and genotype OR was calculated using binary logistic regression, both using SPSS v.16. * Adjusted for, smoking habit, and the risk factors including chronic hypertension, diabetic status, chronic renal disease, and abnormal uterine artery Doppler. Allelic OR was calculated using UNPHASED software release 3.1.5.

Locus	Genotype			Allele	Genotype unadjusted OR (95%CI) P value	Genotype adjusted OR* (95%CI) P value	Unadjusted allelic OR (95% CI) P value
rs4444903	**AA**	**AG**	**GG**	**G**			
FGR (N=243)	9.3	5.6	5.1	**0.38**	0.84 (0.68-1.05) 0.13	0.85 (0.68-1.06) 0.15	0.84 (0.68- 1.05) 0.13
Healthy pregnancy (N=570)	33.5	49.2	17.3	**0.42**			
rs2237051	**GG**	**GA**	**AA**	**A**			
FGR (N=243)	40.0	45.0	15.0	**0.38**	0.92 (0.74-1.14) 0.46	0.93 (0.74-1.15) 0.49	0.92 (0.74-1.15) 0.45
Healthy pregnancy (N=570)	37.3	46.4	16.3	**0.40**			

Table 3-4: Haplotype frequency of the EGF variants in preeclampsia and fetal growth restriction cases and controls.

r^2: linkage disequilibrium between the two alleles in control group. Haplotype with frequency <5% were pooled as rare haplotypes. P value was using likelihood ratio Chi-square. UNPHASED V3.1.5 was used for this analysis.

rs 4444903 (A/G)	rs2237051 (G/A)	r^2	Preeclampsia		FGR	
			Cases	Controls	Cases	Controls
A	G	0.81	60.16	57.23	59.77	56.89
G	A	0.81	34.58	38.14	35.08	38.24
Rare haplotypes			5.26	4.63	5.15	4.87
P value			**0.11**		**0.24**	

3.3.3.3 Association of EGF variants with maternal blood pressure

In the control group, the GG genotype at locus rs4444903 was associated with the highest systolic blood pressure but not diastolic blood pressure (P= 0.04 and 0.26 for systolic and diastolic blood pressure respectively; Table 3-5). Systolic blood pressure (SBP) was (Mean 124.2 mm Hg, SD 17.8) and diastolic blood pressure (DBP) was (Mean 74.0 mm Hg, SD 13.0) for GG compared to the (SBP 119.7 mm Hg, SD 13.9; DBP 71.9 mm Hg, SD 12.0) and (SBP 119.0 mm Hg, SD 15.2; DBP 71.2 mm Hg, SD 11.3) for AA and AG genotypes respectively.

The AA genotype at locus rs2237051 was not associated with either systolic or diastolic blood pressure in the control group. Systolic blood pressure for GG, GA, and AA genotypes respectively was 120.2 mm Hg, SD 14.7; 118.6 mm Hg, SD 14.7; and 123.5 mm Hg, SD 17.9 (P = 0.25) and diastolic blood pressure was

72.0 mm Hg, SD 12.2; 70.8 mm Hg SD 10.9; and 74.2 mm Hg, SD 13.3 (P = 0.35) (Table 3-5).

3.3.3.4 Association of EGF variants with offspring birth weight

Maternal homozygosity for the minor allele at both loci was associated with the lowest birth weight, but this was not statistically significant. Birth weight was 3442 g, SD 635; 3499 g, SD 579; and 3400 g, SD 572 respectively in mothers with AA, AG and GG genotypes at locus rs4444903 (P= 0.31). At locus rs2237051 birth weight was 3453 g, SD 615; 3491 g, SD 592; and 3396 g, SD 590 for genotypes GG, GA, and AA respectively (P = 0.71), (Table 3-5).

Table 3-5: Maternal systolic and diastolic blood pressure and fetal birth weight as a function of the genotype of EGF variants in the control group.

Data represent mean (SD) of the trait investigated. The regression coefficient beta, (95% CI), and P values from linear regression are shown. * Adjustment for age, parity, and smoking habits did not alter the findings for maternal blood pressure, and adjustment for smoking habits did not alter the findings for birth weight (data not shown).

Locus	Systolic blood pressure*(mm Hg) (N=673)	Diastolic blood* pressure (mm Hg) (673)	Fetal birth weight* (g) (N=570)
rs4444903			
AA	119.7 (13.9)	71.9 (12.0)	3441 (635)
AG	119.0 (15.2)	71.2 (11.3)	3499 (579)
GG	124.2 (17.8)	74.0 (13.0)	3400 (572)
Beta(95% I)	1.7 (0.55 to 3.44)	0.75 (-0.56 to 2.6)	-8.30 (-79.56 to 62.96)
P	**0.04**	**0.26**	**0.82**
rs2237051			
GG	120.1 (14.7)	72.0 (12.2)	3453 (615)
GA	118.6 (14.7)	70.8 (10.9)	3491 (592)
AA	123.5 (17.9)	74.2 (13.3)	3396 (590)
Beta(95% I)	0.99 (-0.68 to 2.68)	0.62 (-0.69 to 1.92)	-15.42 (-86.68 to 55.85)
P	**0.25**	**0.35**	**0. 67**

3.3.4 Discussion

In this study of women of white western European descent, the results reported here provide no evidence that the EGF genotype GG at marker rs4444903 confer protection against preeclampsia (Table 3-2). Homozygosity for the A allele at marker 2237051, which is in LD with the G allele at rs4444903 (r^2=0.81), showed the same effect, both fell short of statistical significance. Dissanayake et al (2009) reported the haplotype formed of the G allele of rs4444903 and the G allele of marker rs2237051 as a risk factor for preeclampsia in Sinhalese. The authors suggested that this association is mainly related to the G allele of rs2237051, the minor allele in Sinhalese (Dissanayake et al., 2009). It is difficult to compare results from different ethnic groups when the MAF is variable. It is noteworthy that small effects at or below 10-15 %, outside the margins of the effect detectable by this sample size with adequate statistical power, cannot be excluded. Therefore, results should be interpreted with caution, and confirmation in a larger sample is mandatory.

Interestingly, the maternal genotype at rs4444903 was significantly associated with the highest systolic blood pressure measured during pregnancy in the control group of pregnant women (Table 3-4). Previous studies have indicated that higher EGF levels are correlated with diastolic blood pressure in humans (Lundstam et al., 2007) and with hypertension in experimental animals (Florian and Watts, 1999). Several studies have demonstrated the effect of EGF on vascular smooth muscle contraction through MAPK signalling, increase in intracellular calcium, and prostaglandin production. EGF also potentiates the vasoconstrictor effect of angiotensin II and vasopressin on the vasculature. So

whilst EGF acting locally in the placenta may lowers the risk of preeclampsia, maternal systemic EGF may lead to higher blood pressure in healthy pregnancies. It is noteworthy that previous reports suggested EGF levels in maternal blood and urine are not correlated with EGF levels in amniotic fluid, indicating different sources for systemic maternal EGF and EGF in the fetoplacental compartment (Hofmann and Abramowicz, 1990).

Maternal genotypes at rs4444903 and rs2237051 of the EGF gene were not associated with fetal growth restriction, consistent with the findings of Dissanayake et al (2007). That study also reported a significant association between maternal genotype at these loci and birth weight in healthy pregnancies in Caucasian and Sinhalese women. The present study also found that women homozygous for rs4444903 allele G or rs2237051 allele A gave birth to babies with lower birth weights, although this was statistically non significant, in spite of the larger numbers genotyped. Note that 90 of the 570 controls in the FGR study were also included in the study by Dissanayake et al (Dissanayake et al., 2007). The failure to replicate the significant genetic association reported by Dissanayake et al. may be related to the greater complexity of genetic and non-genetic determinants of birth weight in the maternal controls reported here, which included women with diverse medical and obstetric histories.

These findings add to the picture of contradictory data regarding the maternal levels of EGF and fetal birth weight. Moharam et al (1992) found no variation in maternal urinary EGF levels comparing appropriate and small for gestational age pregnancies (Moharam et al., 1992). In contrast Lindqvist et al (1999)

demonstrated that EGF levels in maternal urine were lower in FGR cases compared to controls (Lindqvist et al., 1999). None of these studies considered the genetic background that might alter EGF expression. Besides favourably affecting the placentation process, a protective effect of a maternal EGF genotype reportedly associated with higher EGF production can be explained by the enhancement of metabolite transfer across the placenta. In an animal model, EGF deficiency induced by maternal sialoadenectomy caused intrauterine growth retardation. Lower maternal EGF levels adversely affected fetal growth through decreasing the expression of placental GLUT3 mRNA which leads to reduced transplacental transfer of glucose and lower glucose concentration in fetal plasma. These effects were corrected by EGF supplementation (Kamei et al., 1999). Furthermore, in humans EGF levels in maternal and cord serum were higher in diabetic pregnancies characterized by high birth weight (Loukovaara et al., 2004, Grissa et al., 2010). However in a transgenic mouse model, offspring over-expressing human EGF were growth retarded, an affect attributed to a reduction in the levels of insulin like growth factor binding protein-3 (IGFBP-3) (Chan and Wong, 2000). It is therefore possible that the association of a maternal EGF genotype characterized by higher EGF production with low birth weight is mediated by the transfer of high-expression alleles to the fetus. Interestingly, the study by Dissanayake et al (2007), which included 105 mother-father-baby trios from UK pregnancies affected by FGR, showed that the fetal G-A haplotype composed of the minor alleles of the two loci was transmitted more frequently from heterozygous parents to their FGR offspring ($P = 0.02$), supporting a role for fetal EGF genotype in the pathogenesis of low birth weight (Dissanayake et al., 2007). This also suggests differences in the genetic and pathophysiological

backgrounds controlling variability of birth weight in normal pregnancies and growth restricted pregnancies.

Previous reports indicated that the G allele of rs4444903 is associated with higher EGF production (Shahbazi et al., 2002, Bhowmick et al., 2004, Vauleon et al., 2007). From biological point of view, maternal EGF locally present in the placenta in early pregnancy could be favourable for the mother and the fetus and protect against preeclampsia by enhancing trophoblast proliferation, differentiation, and invasion and acting as an antiapoptotic factor for the trophoblast cell lineage, so helping in the establishment and maintenance of the placenta. However the current genetic association study did not provide statistically significant support for the protective effect of this variant of EGF against preeclampsia or FGR.

3.4 The effect of c.2556G>A (rs2237051) on EGF mRNA splicing

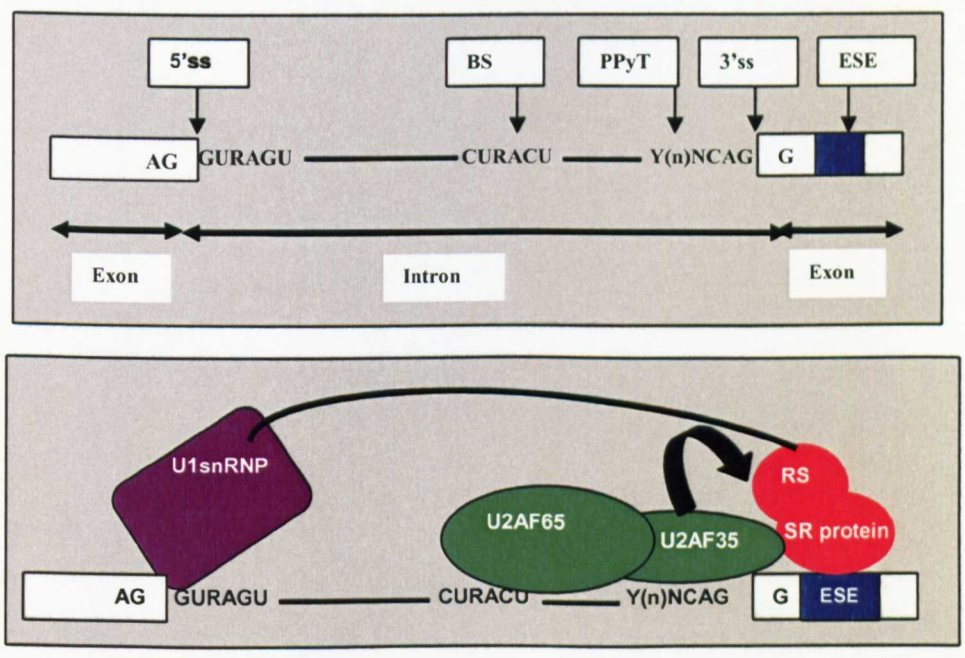
3.4.1 Introduction

Pre-mRNA splicing is a critical step in gene expression in which intron sequences are removed and exons are aligned together to generate a mature mRNA transcript (House and Lynch, 2008) legible for translation. Exon recognition by assembly of large ribonucleoprotein complexes known as spliceosomes is the key step in the splicing process. The formation of the spliceosome involves the step-wise assembly of four small nuclear ribonucleoproteins (snRNPs U1, U2, U4/U6 and U5) and many non-snRNP splicing factors on a pre-mRNA (Blencowe, 2000). The proteins of the spliceosome bind initially to the core splicing signals which include the donor 3' splice site (3' SS) and acceptor (5' SS) splice sites, the polypyrimidine tract (PPyT), and the branch point (BPS) (Figure 3-2). These core splicing signals are found to carry only a small amount of the information required for exon/intron definition (Wang et al., 2008), and proper exon definition requires additional information nested in cis acting elements called splicing regulatory elements (SREs). SREs are sequence motifs (6-8 nucleotides) identified both in introns and exons which function by either facilitating (enhancers) or suppressing (silencers) splicing. Exonic splicing enhancers (ESE) have been largely characterized and reported to be present in most, if not all, mammalian exons (Cartegni et al., 2002, Fairbrother et al., 2002). Purine-rich ESEs promote the recognition of exons through binding of SR proteins (serine–arginine rich proteins) and SR related proteins via their N-terminal RNA Recognition Motif (RRM), and thus facilitating the recruitment of the spliceosome to the vicinity of

exon-intron junctions by protein-protein interaction through C-terminal arginine serine rich domain (RS domain) (Shen and Green, 2006).

Figure 3-2: Core and cis- regulatory splicing elements

Upper figure: Pre-mRNA sequences important for splicing. Consensus mammalian 5′ splice site (5′ SS), branch site (BS), poly-pyrimidine tract (PPyT) and 3′SS sequences are shown. ESE, exonic splicing enhancer; R, purine; Y, pyrimidine, adapted from (Blencowe, 2000). Lower figure: A proposed mechanism of the action of ESE and SR protein in mRNA splicing, SR protein binds to an exonic splicing enhancer (ESE) through its RNA-recognition motifs (RRM) and contacts the various splicing factors at the adjacent splice sites through its RS domain. The three sets of splicing-factor–pre-mRNA interactions (U2AF-3'splice site, U1snRNP-5'splice site and SR protein–ESE) are strengthened by the protein–protein interactions that are mediated by the RS domain, adapted from (Cartegni et al., 2002).



3.4.1.1 Role of mRNA splicing in human diseases

There is a growing body of evidence that ESE disruption by deletion, mis-sense, nonsense or even translationally silent mutations and the use of cryptic splice sites are associated with many diseases and phenotypic abnormalities due to exon skipping or aberrant exon inclusion. For example a C->T mutation at position 6 in exon 7 of survival of motor neuron 2 (SMN2) gene causes skipping of exon 7 due to disruption of an ESE (Cartegni and Krainer, 2002) or creation of an exon splicing silencer (ESS) (Kashima et al., 2007). The protein product of the mRNA lacking exon 7 is unstable and leads to spinal muscular dystrophy. Another example is the microtubule associated protein tau gene encoding the tau protein. Multiple mutations of exon 10 result in inclusion or exclusion of exon 10 with the production of abnormal protein that is easily precipitated in neurofibrillary tangles, the hallmarks of several neurodegenerative diseases such as Alzheimer's disease (Liu and Gong, 2008). Disruptions of functional ESEs are also the principal causes of exon skipping in neurofibromatosis type 1 (Zatkova et al., 2004).

Previous studies in our laboratory showed statistically significant variations in birth weight as a function of genotype at this locus in 3 independent collections of samples from UK Caucasians and Sinhalese from Sri Lanka (Dissanayake et al., 2007). They also showed a distortion in transmission of the G-A haplotype formed of the minor alleles of the rs4444903and rs2237051 to growth restricted babies. The G allele (minor allele in Sinhalese) of SNP rs2237051 is also implicated in the risk of preeclampsia. The GA and GG genotype of this marker was associated with a higher risk of preeclampsia in Sinhalese women compared

to the AA genotype (OR 1.7 (95% CI 1-2.8) and OR 1.8, (95% CI 1-3.2) respectively (Dissanayake et al., 2009).

The current study focuses on the functional analysis and elucidation of the mechanisms underlying this association. One of the proposed mechanisms is that rs2237051 (G>A) causes exon 14 skipping due to disruption of an ESE motif.

3.4.1.2 Analysis of splicing regulatory elements

Analysis of exonic sequence for the presence of putative ESE motifs usually starts using *in silico* analysis tools such as ESE-Finder and RESCUE-ESE to identify putative elements (Cartegni et al., 2003, Fairbrother et al., 2004). ESE-Finder is a program based on the sequence from a library of oligonucleotides that showed favourable splicing after multiple rounds of a splicing assay. The sequences were used to produce scoring matrices for splicing regulatory proteins ASF/SF2, SC35, SRp40, or SRp55 (Liu et al., 1998, Cartegni et al., 2003). RESCUE-ESE motifs were predicted by selecting hexamers that were enriched in exons versus introns and in exons with weak versus strong splice sites.

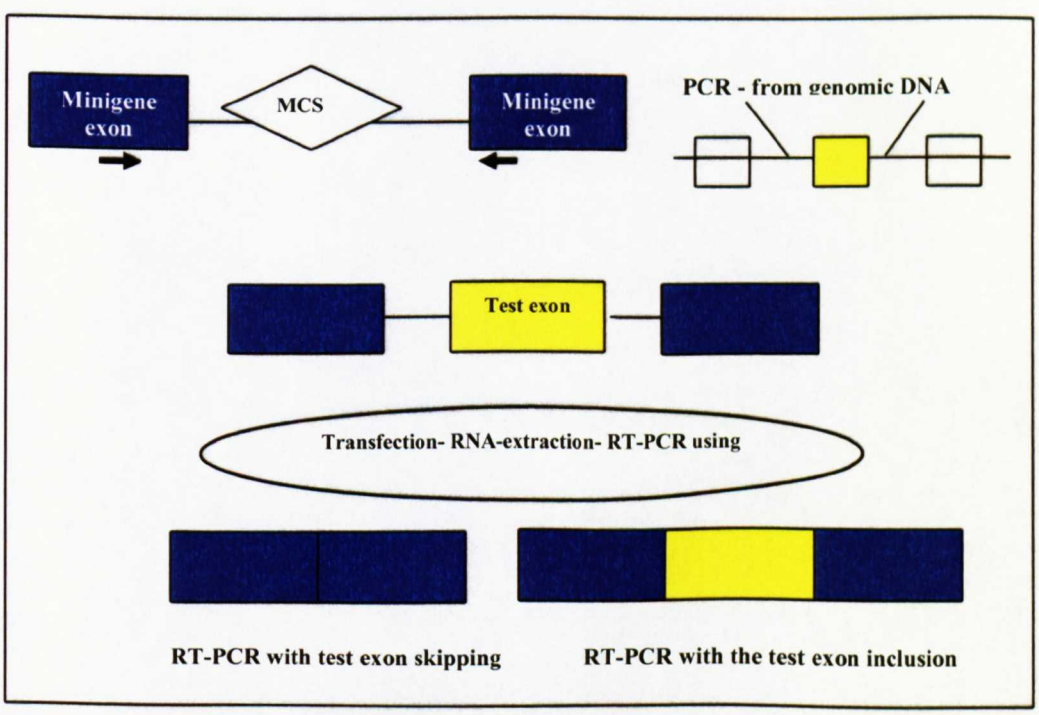
However, the presence of a high-score motif in a sequence does not necessarily indicate that the sequence acts as a functional ESE. Functional assessment of the identified motifs can be carried out by using a minigene technique.

A minigene is a vector that contains at least 2 exons with a known splicing pattern flanking a multiple cloning site for insertion of the genomic segment from the gene of interest. The insert represents the exon carrying the ESE motif

to be investigated and the flanking intronic sequences that carry the primary splicing signals (usually 200-300 of the adjacent introns) amplified from the genomic DNA by PCR. Transient transfection, RNA extraction, and RT-PCR (reverse transcriptase-polymeras chain reaction), using vector specific primers is performed (Cooper, 2005). If a given exon carries a mutation that might affect splicing by disrupting an ESE sequence, 2 minigene constructs with each allele sequence should be tested. If the motif functions as an ESE the PCR product will include both the vector specific exons and the test exon. If the mutation disrupts that ESE motif the PCR product will include vector specific exons only (Figure 3-3).

Figure 3-3: Outline of the minigene technique

Blue boxes are the vector exons, MCS is the multiple cloning site and the black arrows represent minigene specific primers. Genomic DNA is represented by the white boxes (exons) and the connecting line (introns), the yellow box is the test exon, SRE: splicing regulatory element.



3.4.2 Methods

3.4.2.1 EGF exon 14 sequence bioinformatic analysis

Two programs designed to find exonic splicing enhancers were used to examine EGF exon 14: ESE-Finder and RESCUE-ESE. ESE-Finder (http://rulai.cshl. edu/tools /ESE/) is a web-based resource that facilitates rapid analysis of exon sequences to identify binding motifs for four SR proteins: SF2/ASF, SC35, SRp40 and SRp55.20. The RESCUE-ESE program (http://genes.mit.edu /burgelab/rescue-ese/) uses a computational method that identifies exonic splicing enhancers in human genomic sequences by searching for specific hexanucleotides. Further sequence analysis was done to test for the strength of the splice site (http://www.fruitfly.org/seq_tools/splice.html) and other SR and SR related protein binding sequence as well as ESE and ESS motifs using the Splicing Sequences Finder (http://www.umd.be/SSF). The default thresholds of these programmes were used as reference values.

3.4.2.2 Minigene constructs

EGF exon 14 and flanking intronic sequence representing major and minor alleles respectively (rs2237051 G and rs2237051 A) were derived by PCR amplification of genomic DNA (gDNA) from homozygous subjects. The primers used were: EGF_E14S, 5'-TCCCATGTCGACGTGACGTTCCCTCCTCATGT-3' and EGF_E14AS, 5'-GGTGAATCTAGAGGGATAAAGCCCCAATTACC-3' forward and reverse primers respectively. These primers introduced Sal I and Xba I restriction sites (underlined) in the 5' and 3' ends of the fragment respectively, to be used for further cloning. The PCR reaction was performed in a 30µl reaction volume containing 200ng of gDNA, 1XTaq buffer containing 1.5 mM $MgCl_2$, 200 µM dNTPs, 0.5 µM of each primer, and 1 unit of Taq

polymerase. The conditions were 94°C for 40 s, and 35 cycles of denaturing at 94°C for 30 s, primer annealing at 60°C for 1 min, and extension at 72°C for 1 min, with a final extension of 10 min at 72°C. PCR products were visualized by electrophoresis on 1% agarose gels stained with ethidium bromide. The PCR fragment was 574 bp long (202 bp of intron 13, 168 bp of exon 14, and 204 bp of intron 14), ensuring that all splicing signals for exon 14 were included. The PCR products were TA cloned into a TOPO2.1 vector followed by transformation of One Shot® TOP10 Chemically Competent E. coli according to the manufacturer's protocol (Invitrogen).

After DNA mini-preparation (QIAGEN), inserts were sequenced. The reactions were carried out in 10µl using 5pmol M13 reverse primer 5'-CAGGAAACAGCTATGAC-3' and 1x Big-dye 3.1 reaction mix, 1x Big Dye terminator buffer (Applied Biosystems) and 400 ng DNA. These underwent 25 cycles on a thermo-cycler as follows: denaturation at 96°C for 30 s, primer annealing at 50°C for 15 sec and extension at 60°C for 4 min. The reaction products were purified as described earlier (Section 2.4.5.5). Sequencing was performed using an ABI-automated sequencer. Sequencing alignment was done using EBI-tools- ClustalW web-based software (http://www.ebi.ac.uk/Tools/clustalw2/index.html) to confirm that the inserts were identical apart from the base at rs2237051 (G>A).

Restriction digestion of the TA constructs and destination vector (RCHGLO minigene kindly provided by Prof. T. Cooper (Ray et al., 2006) was carried out in 10µl reaction volumes containing 5 units each of Sal I and XbaI, 1x digestion

buffer, and 2 μg DNA for 2 hours at 37°C. The reaction was heat inactivated at 65°C for 10 min and the fragments were gel extracted using gel extraction kits (QIAGEN). 100 ng of the vector and 30ng of the insert (1:3 ratio) were subjected to a ligation reaction using 1U T4 DNA ligase, 1x ligation buffer in 10μl at 4°C overnight. Transformation of X1-blue super-competent cells was carried out according to the manufacturer's protocol (Stratagene). The colonies containing the recombinant vector were selected by restriction enzyme treatment using Sal I and Xba I, and were sequenced as above using vector primer TNIE4: 5'-AGGTGCTGCCGCCGGGCGGTGGCTG-3' and the insert primer EGF_E14AS: 5'-GGTGAATCTAGAGGGATAAAGCCCCAATTACC-3. For the RCHGLO constructs containing the correct sequence (RCHGLO-EGF-G and RCHGLO-EGF-A), a large plasmid preparation was made using Plasmid Purification endotoxin-free Midi Kit (QIAGEN).

3.4.2.3 Cell culture and transfection

HepG2 (hepatoma cell line) cells were maintained in Eagle's minimal essential medium (EMEM) with 10% FCS, 10 mM L-glutamine, penicillin (100U/ml), streptomycin (100 pg/ml), Amphotericin B (2.5 μg/ml) at 37°C and 5% CO_2 in 75 ml culture flasks. SGH-PL4, (extravillous trophoblast cell line) cells were maintained in F-10 Ham Nutrient Mixture with L-glutamine and the same antibiotic and culture conditions as HepG2. The cells were subcultured once they approached 80% confluence (usually after ~3 days of growth). Cells in a single 75 cm^2 flask were split equally into 3 new flasks as part of this process. The passage number was limited to 8-10 times, over the experiment after which fresh aliquots were taken from the liquid nitrogen to establish consistency in the

growth pattern and cellular characteristics over the experiments. For transfection 2×10^6 cells were plated in 60mm culture plates 24 hours before transfection. Transfections were performed using 1μg plasmid DNA and TFx20 transfection reagent (Promega) in 3:1 charge ratio. The DNA:TFx20 mixture was incubated at room temprature in 2 ml serum free medium for 15min then transferred to cells and incubated for 1 hour at 37°C after which 3 ml of complete medium was added. Because the assay applied in this experiment was a qualitative assessment of splicing pattern using vector specific primers for RT-PCR to confirm that transfection was successful, we did not opt to control for transfection efficiency. However if different transcripts were noticed and comparison between the levels of these transcripts was required controlling for transfection efficiency would be important. Co-transfection by vector expressing fluorescent protein like green fluorescent protein that can be visualized would be an option, or the use of flowcytmetry sorting, an expensive option can be used to control for transfection efficiency. The number of dead cells was trivial and comparable in all plates. Twenty four hours later, cells were harvested and prepared for total RNA extraction. All transfections were performed in three independent experiments.

3.4.2.4 RNA extraction and RT-PCR

Total RNA was isolated from HepG2 and SGH-PL4 cells using RNeasy kit (QIAGEN) with application of the DNase treatment step to ensure the removal of any genomic and plasmid DNA. Reverse transcription was performed in 20μL reactions using 2μg of total RNA, 300 ng random primers (Promega), plus RNase free water. This reaction was heated at 65°C for 5 min and then chilled on ice. 1X Superscript-RT buffer, 200mM DTT (Dithiothreitol), 80mM dNTPs, 1U

RNase Block and 2U of Superscript-RT (Invitrogen) was then added to the reaction and heated as follows: 25°C for 10 min, 42°C for 60 min, 70°C for 15 min and 10°C hold. As a control for contamination with plasmid DNA, duplicate samples were used in which reverse transcriptase was not added to the mixture in one sample. For the RT-PCR the following pair of vector specific primers was used: sense primer RSV5U: 5'-CATTCACCACATTGGTGTGC-3' and the antisense primer TNIE4: 5'-AGGTGCTGCCGCCGGGCGGTGGCTG-3'. The amplification was carried out in 30μL reactions, containing 1μL of cDNA, 1XTaq buffer with $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 200 μM dNTPs, 0.5 mM of each primer, and 1 unit of Taq polymerase. The PCR conditions and analysis were the same as described above.

3.4.3 Results

3.4.3.1 Initial in silico sequence analysis

ESE-Finder showed that EGF exon 14 contains a putative SF2/ASF protein binding site ESE sequence motif (GAGAGTA) within the rs2237051 G allele with the potential to enhance splicing and inclusion of exon 14 in the mature mRNA. The rs2237051 G>A SNP (AAGAGTA) disrupts this site, potentially leading to loss of enhancer function with possible exon 14 skipping (Figure 3-4) The output of RESCUE-ESE also showed that substitution of A for G completely eliminates the hexamer ATGAGA (Table 3-6).

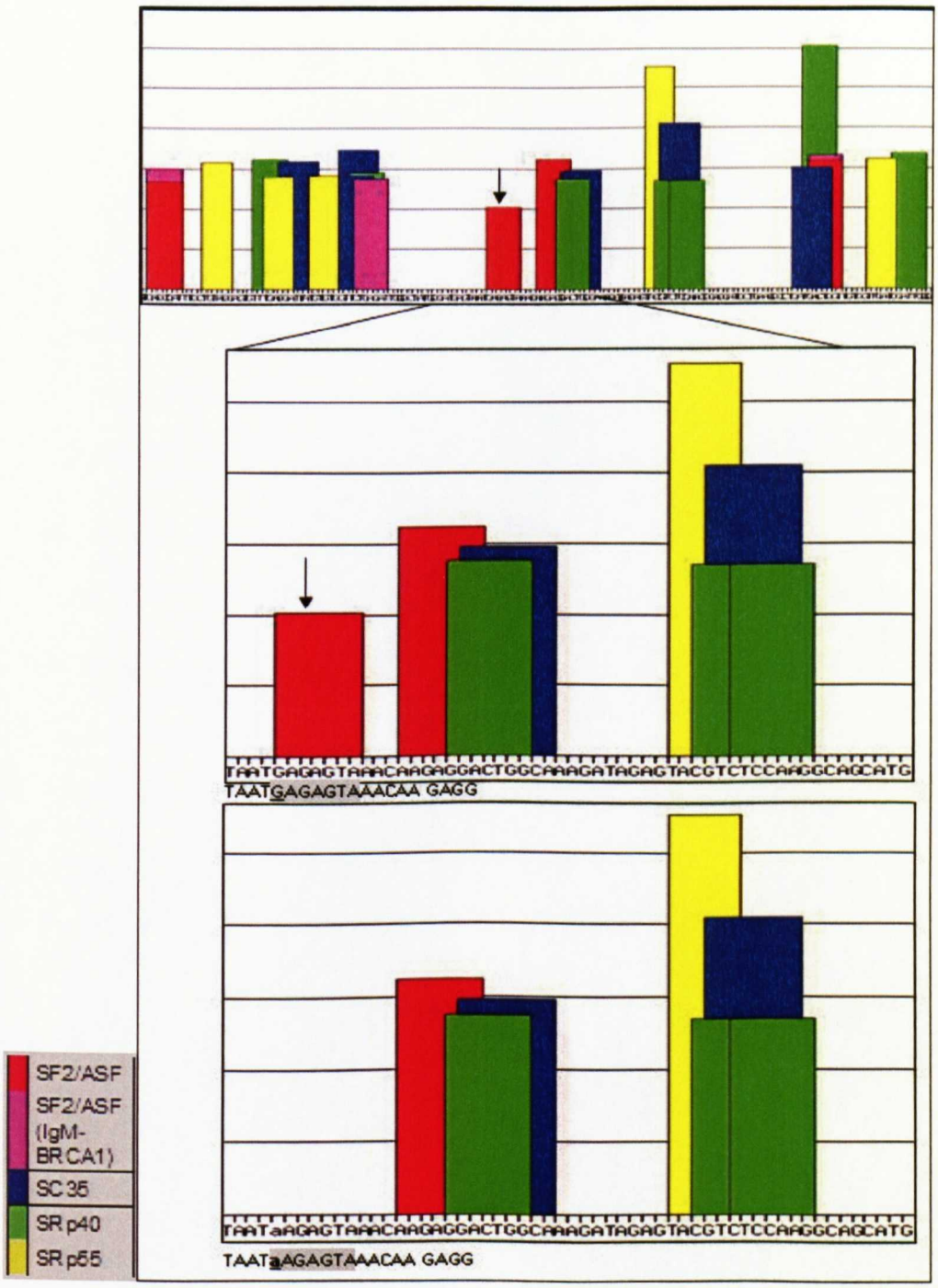
Exon 14 is a well defined exon considering the strength of both acceptor (score 0.87) and donor site (score 0.98) and the consensus 3'and 5' SS. Although the polypyrimidine tract of the preceding intron has no more than 3 uridine stretches, based on the analysis of the sequence -4/-25 (tttaacaaacttgaattgtttc) a high pyrimidine:purine ratio (13:9) was found just upstream to the 3' SS, making that PPyT potentially effective in splicing.

Further exploration of the G and A alleles of exon 14 sequence using http://www.umd.be/SSF, showed that while rs2237051 G>A destroyed the low threshold SF2/ASF and SF2/ASF(IgM-BRCA1) ESE motif (GAGAGTA), it also creates another 3 SR protein binding ESEs, 2 with higher thresholds for TRA2-B and one for 9G8. It also disrupts a weak ESS hnRNP (heterogeneous nuclear ribonucleoprotein protein) A1 binding site, introduces a moderately powerful one, and reduces the threshold for a third one (Table 3-6).

In total 25 SR protein binding ESE motifs and 17 hexamer sequences that may act as ESEs were identified using bioinformatic tools in exon 14. It is noteworthy that 2 of the 25 ESE motifs have a very strong threshold, indicating that these 2 motives might be the functional ones in EGF exon 14 (Figure 3-4).

Figure 3-4: ESE-Finder graphical output for exon 14

Upper panel: full length exon 14; middle panel: the G allele; lower panel: the A allele. The arrow refers to the ESE-SF2/ASF protein binding motif in the upper and middle panels; this motif is lost in the A allele shown in the lower panel. The highlighted text is the protein binding motif and the SNP is underlined. The height of the bars indicates higher threshold for binding proteins.



3.4.3.2 In vivo splicing assay

To validate these findings, hybrid minigene constructs from both major (RCHGLO-EGF-G) and minor alleles (RCHGLO-EGF-A) of exon 14 were prepared. The insert included both exon 14 and the flanking intronic sequence which contains the core splicing signals (5' SS, BPS, PPyT and the 3' SS) so that any splicing difference between G and A alleles is due to SNP rs2237051 G>A (Figure 3-5, A). The *in vivo* splicing experiments were performed by transfection of the minigene constructs into HepG2 (hepatoma) and SGH-PL4 (extravillous trophoblast) cell lines to study the splicing reaction in two cellular contexts and exclude any splicing variation due to tissue specific changes in the trans-acting splicing proteins. The expected size of the spliced product was 248bp representing 40bp from each minigene exon with 168 bp of exon 14 included. The expected size of the unspliced product was 80 bp due to exon 14 skipping. Identical patterns of splicing were observed for G and A alleles in both cell lines, generating 248 bp transcripts with inclusion of exon 14 (Figure 3-5, B).

Figure 3-5: Minigene constructs and *in vivo* splicing assay

A (a): Schematic presentation of exon 14 showing 3'ss, 5'ss and the G allele ESE motif. (b): Schematic presentation of the minigene construct. The red box is exon 14; white boxes are the vector specific exons (VSE); arrows indicate the vector specific primers used in the RT-PCR. (c): the expected 248bp PCR product if the mRNA contains exon 14, and the expected 80bp PCR product if exon 14 has been skipped. (B) RT–PCR analysis of EGF exon 14 after transient transfection of SGH-PL4 cell line by hybrid minigene-EGF exon 14 G and A alleles (representative of triplicate independent experiments in both HepG2 and SGH-PL4 cell lines) in which both the G and A alleles show an identical pattern of splicing with inclusion of exon 14.

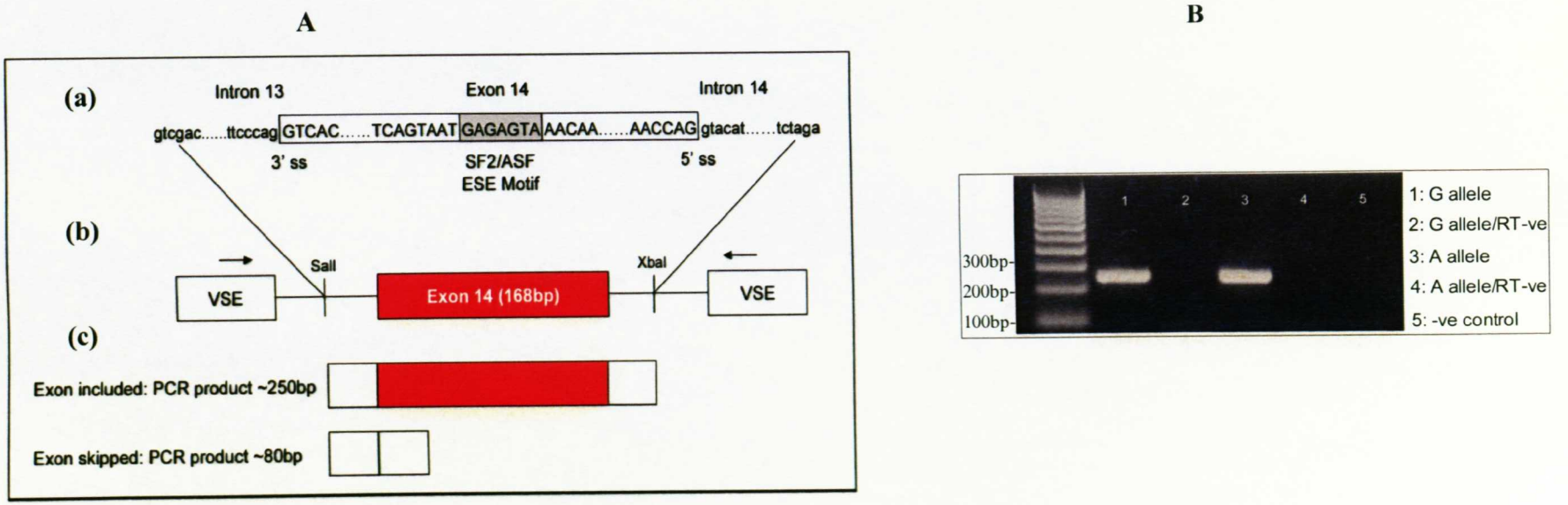


Table 3-6: Tabular presentation showing the effect of rs2237051 G>A on different cis acting regulatory elements in the vicinity of the polymorphism

G allele sequence: TCAGTAATGAGAGTAAACAA. A allele sequence: TCAGTAATAAGAGTAAACA. *The SNP is shown in bold in major and minor alleles. **Threshold values for the different binding proteins are as follow: SF2/ASF: 72.98; SF2/ASF (IgM-BRCA1): 70.51; Tra2: 75.964; 9G8: 59.24; hnRNP A1: 65.476, and silencer motif by Sironi et al: 60.

A: Enhancer motifs						
Software	Protein linked	Major allele motif*	Score**	Minor allele motif	Score	Effect
ESE-Finder matrices	SF2/ASF (IgM-BRCA1)	GAGAGTA	(70.92)			Site broken
	SF2/ASF	GAGAGTA	(73.44)			Site broken
ESE motifs from SSF	Tra2-β			AATAA	(98.53)	New site
	9G8			TAAGAG	(63.16)	New site
	Tra2-β			AAGAG	(89.07)	New site
Rescue ESE hexamer		ATGAGA				Not found
B: Silencer motifs.						
hnRNP motifs from SSF	hnRNP A1	TAATGA	(67.62)			Site broken
	hnRNP A1	GAGAGT	(76.19)	AAGAGT	(72.38)	-5 %
	hnRNP A1			TAAGAG	(71.67)	New site
Silencer motifs from Sironi et al by SSF		TGAGAGTA	(68.15)	TAAGAGTA	(67.22)	-1.36 %

3.4.4 Discussion

Large numbers of human synonymous and non synonymous mutations have been implicated in disruption of mRNA splicing, linking coding sequence mutations to genetic diseases. Although initial sequence analysis predicted disruption of a putative SF2/SAF binding motif in the A allele (variant allele) of exon 14 of the EGF gene, the *in vivo* splicing assay in 2 different cell lines showed that this alteration is not functionally significant and does not lead to exon 14 skipping.

The lack of functional effects secondary to disruption of this ESE motif raised several possibilities. The predicted G allele motif might be functionally neutral or part of a Composite Regulatory Elements of Splicing (CRES) that functions both as an enhancer and silencer according to their sequence context. The second possibility is that the G allele motif might overlap other ESE motifs not tested for by the programs ESE-Finder and RESCUE. Lastly, the SNP G>A might create a new ESE sequence motif or, it might destroy an ESS sequence motif.

The fact that rs2237051 G>A does not affect splicing of exon 14 might indicate either that exon 14 splicing is not dependent on this motif, or that the tested motif is present by chance, or that it is neutral in the context of EGF exon 14. In agreement with this, it has been suggested that the vast number of ESEs in human exons, and the short and degenerate nature of the ESE elements reduces the possibility that all of them are binding targets of the SR proteins (Wu et al., 2005). For example, the consensus ESE sequence for the splicing regulatory

protein SF2/SAF is CRSMSGW, indicating that the sequence is not exclusively specific as 5 out of 7 nucleotides can vary within this short sequence motif. It is worth noting that the full length exon 14 has at least 24 other SR binding sequences, many of which have higher scores than the tested ESE. It is possible that one of those high score motifs might be a genuine ESE (Figure 3-4) if any is required. The strength of the splice sites, the position of the ESE along the exon, and the length of the flanking intron play a significant role in ESE activity and dependence (Dewey et al., 2006, Hertel, 2008, Cartegni et al., 2002). In this respect, it is worth noting that EGF exon 14 is strongly defined by its splice site and that the ESE element analyzed here is located 70 nucleotides from the 5' splice site and flanked by long introns; all of those reduce the dependence of EGF exon 14 splicing on ESE sequences.

There is a possibility that the tested ESE sequence is a part of a Composite Regulatory Element of Splicing (CRES) with overlapping silencer and enhancer functions that are context dependent rather than SR protein dependent (Pagani et al., 2003) and altering the binding affinity to SR protein does not interfere with its function. It has been suggested that the term enhancer or silencer is appropriate only if the context is considered; an ESE element may well act as a negative regulator of splicing in another context, e.g., when inserted into flanking intron regions or other exons (Holste and Ohler, 2008).

Analysis of the G allele sequence using SSF shows nearly complete overlap between ESE (SF2/ASF; GAGAGTA) and 3 ESS sequences (Sironi silencer

Motif 2: TGAGAGTA (Sironi et al., 2004); hnRNP A1: TAATGA, and hnRNP A1: GAGAGT). Substitution of A for G destroys the SF2/ASF binding ESE but introduces 3 new ESE motifs, 2 for Tra2 B and one for 9G8. Substitution of A for G also affects silencer motifs: one hnRNPA1 ESS binding motif is destroyed, but a new one is introduced (Table 3-6). Considering the competitive ((Expert-Bezancon et al., 2004) and spatial constraints, overlapping binding sites will have positive or negative effects depending on which proteins bind to them, and this plays an important role in the regulation of numerous splicing events. Although this Minigene experiment has not determined which, if any, are functional in the exon 14 environments, theoretical considerations suggest that rs2237051 G>A might increase the overall enhancer activity, as when ESS sequences operate in conjunction with ESE sequences, the latter are dominant over the adjacent ESS sequence (Disset et al., 2006).

There are limitations to the use of sequence analysis alone in identifying functional ESE, as bioinformatic tools tend to over-estimate the ESE. Vast numbers of SRE are currently defined all of which are 6-8 nucleotides in length and highly degenerate in nature. In spite of the loss of sequence context in a minigene assay, this approach provides a better method of assessing loss and gain of function for cis and trans-acting factors that affect splicing until the precise interplay between the core splicing elements and cis and trans-acting SREs in defining the "splicing code" is elucidated (Cooper, 2005).

3.5 Functional Characterization of the EGF 5'UTR genetic variation: translational versus transcriptional effects

3.5.1 Introduction

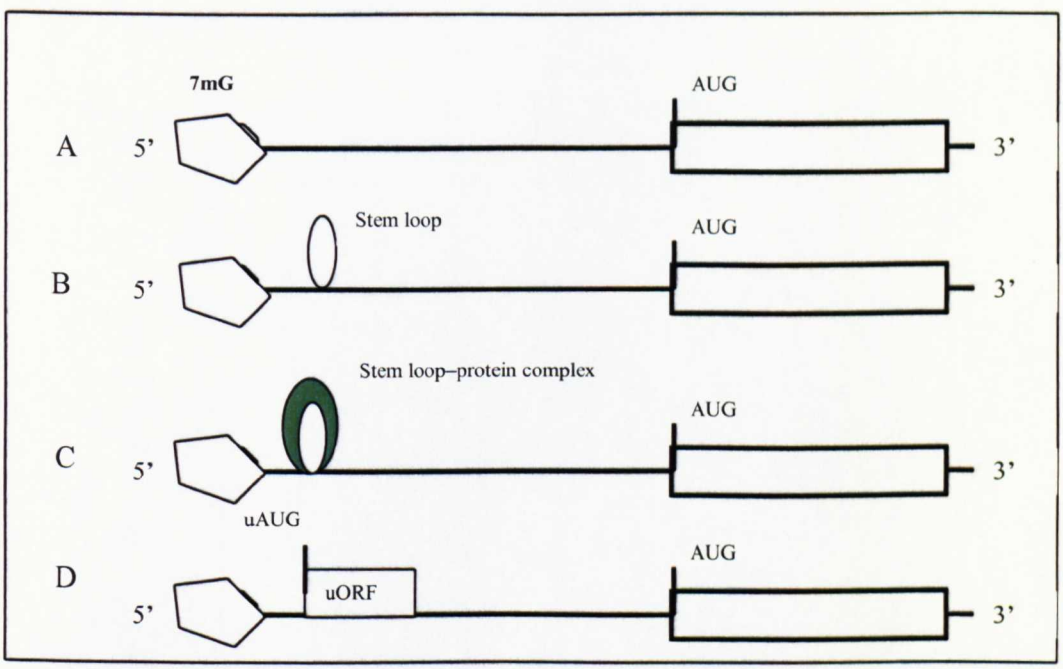
Translation of mRNA into protein is a necessary constitutive step in the gene expression process. Translational control enables a cell to control its protein requirements very rapidly which is of great importance, particularly in the regulation of genes implicated in cell growth and differentiation (Cazzola and Skoda, 2000).

Translational control can be exerted by global machinery that affects all cellular mRNAs by alteration of the available initiation factors and ribosomal proteins (Kozak, 2005). In addition, specific mechanisms that influence the function of particular mRNAs through sequence motifs in the mRNA untranslated regions (UTRs) are also recognized. Many features reside in the 5' UTR and are known to impede the translation of the downstream gene. The presence of stem loops will interfere with the formation of the initiation complex and ribosomal scanning of message and consequently lower the efficiency of translation particularly if the stem loops create structures which are targets for mRNA binding proteins which regulate translation. These stem loops arise due to self-complementary sequence, with formation of stable structured mRNA that requires increased energy for unwinding. The second major feature which alters the efficiency of translation is the presence of upstream open reading frames (uORF) initiated by AUG codons. These uAUGs can cause premature initiation and inhibit translation by preventing the ribosome from reaching the

physiological start codon (Figure 3-6). The efficiency of translation is also reduced in genes in which the 5'UTR is simply longer than the average UTR length of 70-200 base pairs (Kozak, 2005).

Figure 3-6: Features by which 5'UTRs may reduce the efficiency of translation

(A) mRNA with no negative regulatory element, showing the 7 methyl guanine (7mG) for initiation of translation, and the AUG start codon. (B) Self-complementary sequences with stable stem-loop structures (C) stem loop with RNA-binding proteins (green). (D) A 5'-UTR which contains an upstream open reading frame (uORF).



Earlier in this chapter (Section 3.3) we showed that SNP rs4444903 (c.61A>G) of the EGF 5'UTR has a potential protective effect against the risk of preeclampsia. Previous studies in our laboratory showed that the minor G allele was associated with low birth weight and FGR (Dissanayake et al., 2007). In the field of cancer research, a number of investigators have demonstrated that the G allele of c.61A>G is associated with different types of cancer, including

melanoma, glioblastoma, and hepato-cellular carcinoma (Bhowmick et al., 2004, Shahbazi et al., 2002, Tanabe et al., 2008). Others researchers attempting to reproduce these findings failed to show an association with an increase in cancer risk (Randerson-Moor et al., 2004, Vauleon et al., 2007).

To link a genetic variation to a disease entity it is important to show that the gene has a feasible physiological function in the tissue of interest, and that the polymorphism causes a relevant alteration in the level or the function of the gene product. In this regard, EGF is a major regulator of the placentation process (reviewed in Section 3.3.1) and the EGF 5'UTR contains features which may affect the translation of the downstream gene that merit investigation.

The 5'UTR of the EGF is twice as long as an average 5'UTR (442 base pairs compared to 70-200 base pairs average UTR length). It carries 2 upstream open reading frames (uORF) 5' of the ATG which initiates translation of preproEGF. It is predicted to form a structured mRNA. All of these features are reported to lower the efficiency of gene translation in the context of other genes (Chatterjee and Pal, 2009).

3.5.2 Hypothesis and aims

The rs4444903 polymorphism in the EGF-5'UTR is expected to alter the efficiency of translation of the EGF gene and consequently contribute to either increased or decreased gene expression or function. If this is true mRNA levels transcribed from the A and G alleles would be the same whilst levels of the final protein product would differ. This is in contrast to a transcriptional effect in

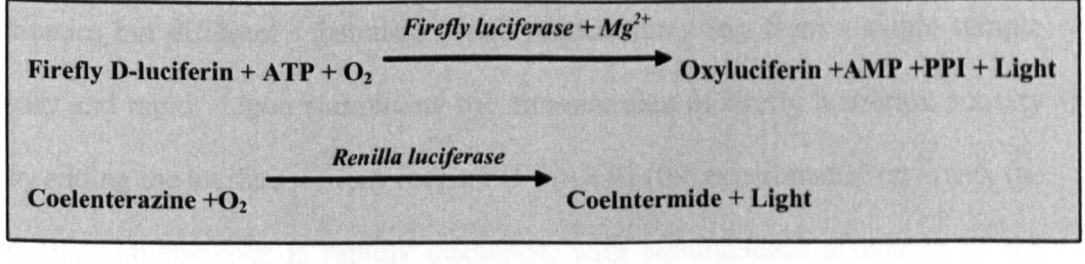
which the allele-specific levels of both the mRNA and protein products would differ. The hypothesis was tested using a luciferase reporter gene assay, and *in vitro* transcription and translation (IVT).

3.5.2.1 Luciferase reporter gene assay

Reporter gene assays are based on the measurement of a reporter gene protein expressed in mammalian cells following transfection with the DNA sequence of interest in order to test its effect on the level of reporter gene expression. This gives an understanding of how this DNA sequence may affect the expression of the native gene *in vivo*. The reporter gene product should not be produced by the mammalian cells, and it should be easy to measure, quantitative, reproducible and sensitive to a wide range of concentrations.

Luciferase refers to a family of enzymes that catalyse the oxidation of various substrates (e.g. luciferin and coelenterazine), resulting in light emission. Firefly luciferase and Renilla luciferase are produced from two distinct organisms, American firefly photinus pyralis and renilla reniformis respectively. They catalyze two different reactions as shown in (Figure 3-7).

Figure 3-7: Firefly and Renilla luciferase mediated reactions



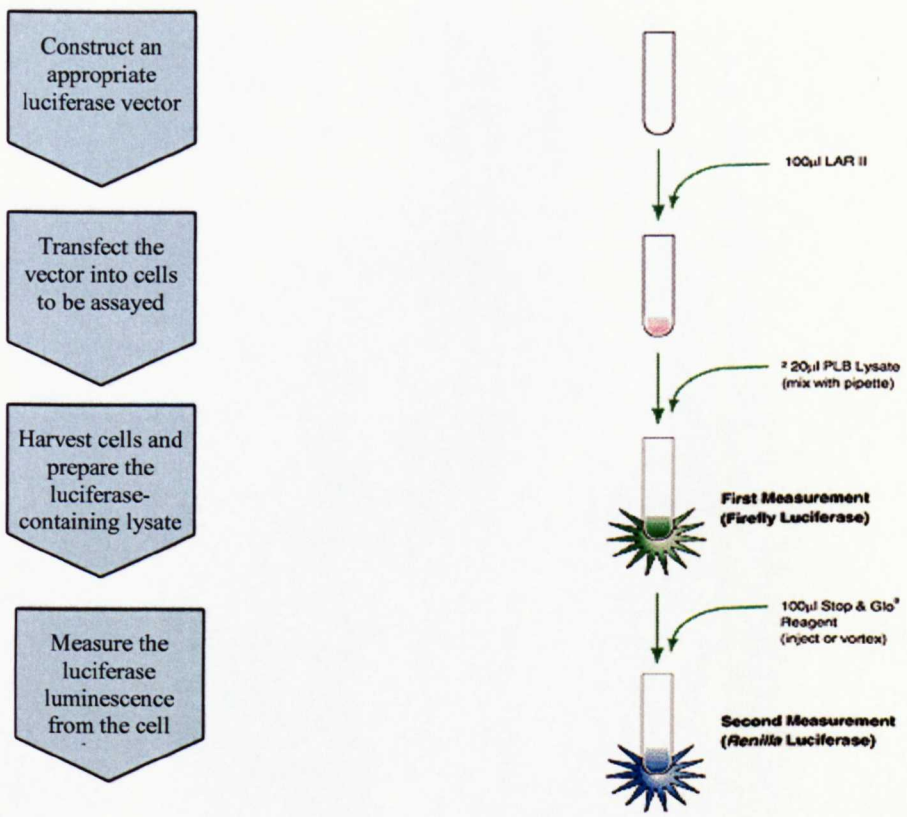
Inevitable experimental variations during the reporter gene assay such as cell number, pipetting volume errors, transfection efficiency, cell lysis efficiency, and assay efficiency can be minimized by using a dual reporter gene assay. This makes use of two different reporters, one as the experimental reporter and the other as a control to normalize the activity of the experimental reporter with respect to experimental errors. The experimental reporter is expressed partially under the control of the test DNA sequence introduced in the plasmid constructs and partially under the control of elements integrated in the reporter gene vector. The control reporter is constitutively expressed, and its activity correlates to the amount of DNA transfected into the cells and the cell's ability to produce protein under the transfection conditions applied. So normalizing the experimental reporter activities to the control reporter activities adjusts for well to well transfection variability and efficiency. Throughout these experiments the maximum acceptable coefficient of variation (CV%) between each experimental triplicate was limited to 20%, improving the statistical power of the study by minimising the background noise from analytical imprecision.

The dual luciferase reaction (DLR) assay system from Promega benefits from the fact that both firefly and renilla luciferase mediate reactions with the same kinetics but different substrates, which renders the assay from a single sample easy and rapid. Upon completing the measurement of firefly luciferase activity by adding the luciferase assay reagent II (LARII) (the experimental reporter), the firefly luminescence is rapidly quenched, with simultaneous activation of the renilla luciferase luminescent reaction (the control reporter activity) using Stop

and Glo reagent. The reaction takes 2-3 seconds to start and the activity is measured over 10 seconds (Figure 3-8).

Figure 3-8: Dual luciferase assay

Left figure: a summary of the steps required to perform a luciferase reporter gene assay. Right figure: measurement of firefly and Renilla luciferase activity in the Promega dual luciferase assay.

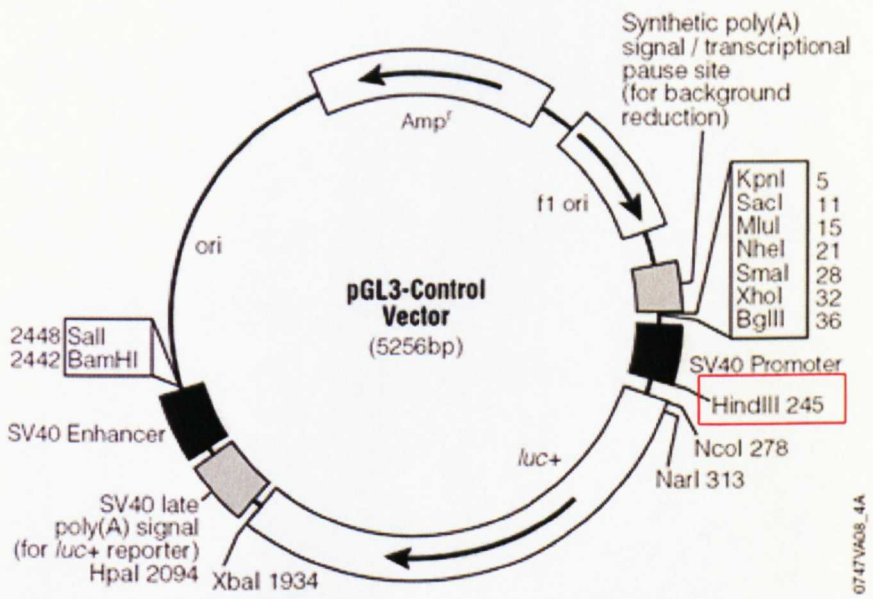


The different expression levels that can be obtained from different pGL3 vectors, pGL3control (pGL3C), pGL3 promoter (pGL3P) and pGL3basic (pGL3B), were exploited in these experiments. pGL3C is a vector which harbours the highly active Simian virus 40 promoter element (SV40P) and SV40 enhancer element (SV40E) upstream and downstream of the luciferase gene (Figure 3-9). pGL3P harbours only SV40P upstream of the luciferase gene.

pGL3B is a promoterless vector which harbours neither of these elements. Due to the higher basal transcription levels from pGL3C and pGL3P vectors, any variation in luciferase output after insertion of the 5'UTR region of EGF is expected to be translational in origin, whilst variation in the luciferase output from the pGL3B is expected to be transcriptional in origin.

Figure 3-9: pGL3 Control vector

The pGL3C vector differs from pGL3 Promoter and pGL3 Basic in that the pGL3P lacks the SV40 enhancer and the pGL3B lacks both the SV40 enhancer and SV40 promoter. The DNA fragment of interest was inserted at the HindIII site (red box) (Promega)



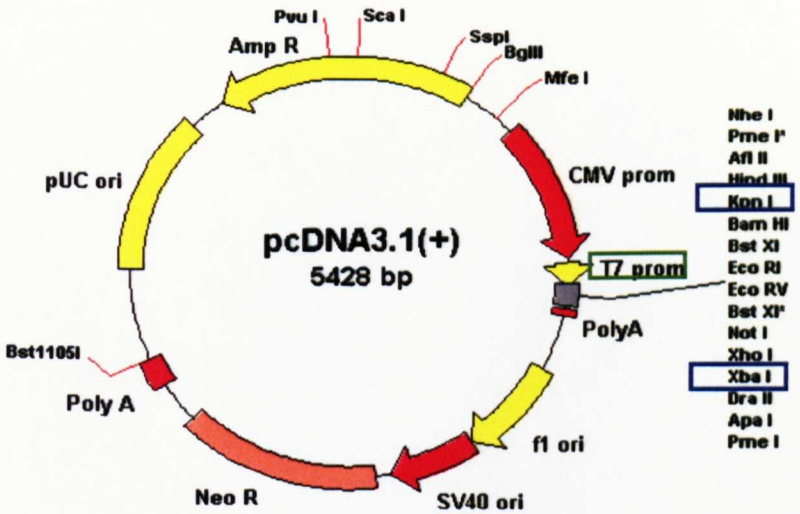
3.5.2.2 In vitro transcription and translation assay (IVT)

In vitro transcription and translation (IVT) methods were used to confirm the effects of the rs4444903 polymorphism on translation in a cell free environment. The principle is to generate a homogenous RNA population representing the 5'UTR of EGF and the luciferase gene after insertion into a pcDNA3.1 expression vector. The construct was inserted downstream of the T7 promoter and the vector linearized just downstream to the insert. T7-RNA polymerase, a bacterial derived enzyme, was used to initiate transcription from its specific promoter sequence (Figure 3-10).

After generating the homogenous RNA from allele-specific constructs, equal amounts were subjected to *in vitro* translation in rabbit reticulocytes. These are prepared from red blood cells after losing their nuclei, but still containing all the molecular components (80S ribosomes, transfer RNAs (tRNAs), aminoacyl-tRNA synthetases, initiation, elongation and termination factors) required for translation of exogenous RNA. For efficient translation, AA, ATP, creatine phosphate, creatine phosphokinase and other co-factors (Mg^{2+}, K^{+}) were added to the reaction. Any variation in protein output from equal amounts of RNA from each allele will be due to variation in translation efficiency.

Figure 3-10: pcDNA3.1 (+) vector

This vector is used for in vitro transcription. DNA fragments were cloned downstream to T7 promoter (green box) after restriction digestion by Kpn1 and Xba1 (blue boxes).



3.5.2.3 Electrophoretic Mobility Shift Assay (EMSA)

The Electrophoretic Mobility Shift Assay (EMSA) is a technique used to characterize protein:DNA interactions. It is based on the fact that complexes of protein and double-stranded DNA migrate through a nondenaturing polyacrylamide gel more slowly than DNA alone. The assay is performed by incubating the protein, or a mixture of proteins, with biotinylated DNA fragments containing the putative protein binding site and nonspecific DNA competitors such as poly (dI-dC). The poly (dI-dC) polymers provide an excess of nonspecific sites to adsorb proteins in a crude nuclear lysate that will bind to any DNA sequence. The complex is then analyzed on polyacrylamide gel. The ability to determine protein:DNA complexes depends largely upon the stability of the complex during its migration into the gel. A low ionic strength electrophoresis buffer is used to increase the stability of the DNA:protein

complex. The specificity of an observed DNA binding reaction can be evaluated using competition assays in which an excess of specific and non specific unlabeled probe are added together with the labelled probe. Specific DNA binding will be eliminated by 100-fold molar excess of unlabeled specific probe, and maintained on addition of unlabeled non specific probe. In addition, specific antibodies can be added to the gel shift reaction. If the protein that forms the complex is recognized by the antibody, the antibody can either block complex formation, or it can form an antibody-protein-DNA complex, resulting in a further reduction in the mobility of the protein-DNA complex (supershift).

3.5.3 Methods

3.5.3.1 Cloning of the 5'UTR of the EGF-luciferase reporter constructs

442 bp of the 5' UTR region of the EGF gene starting from the transcription start site (+1) and extending to the last base before the translation start site was amplified by PCR from genomic DNA previously genotyped for the rs4444903 A>G SNP. 2 DNA samples which were homozygous for A and G alleles respectively were used as template. Amplification was carried out using EGF-UTR-S as forward primer and EGF-UTR-AS as reverse primer (Table 3-7). These primers were flanked on both sides by HindIII restriction sites for use in the cloning procedure.

The PCR reaction was performed in 30µl volumes containing 200ng of gDNA, 1xTaq buffer containing 1.5 mM $MgCl_2$, 200 µM dNTPs, 0.5µM of each primer, and 1 unit of Taq polymerase. The conditions were 94°C for 40 s, and 35 cycles of denaturing at 94°C for 30 s, primer annealing at 61°C for 1 min, and extension at 72°C for 1 min, with a final extension of 10 min at 72°C. PCR products were visualized by electrophoresis on 1% agarose gels stained with ethidium bromide. Fresh PCR products were TA cloned into TOPO.2.1 vector according to the manufacturer's protocol (Invitrogen).

After DNA mini-preparation (Qiagen), inserts were sequenced. The reactions were carried out in 10µl volumes using 5 pmol M13 reverse primer and 1x Big-dye 3.1 reaction mix, 1x Big Dye terminator buffer (Applied Biosystems) and 400 ng DNA. These underwent 25 cycles on a thermo-cycler as follows:

denaturation at 96°C for 30 s, primer annealing at 50°C for 15 sec and extension at 60°C for 4 min. The reaction products were purified using DTR gel filtration cartridges (Edge Biosystems), then heated to dryness at 90°C. Sequencing was performed using an ABI-automated sequencer. Sequencing alignment was performed using EBI-tools ClustalW web-based software (http://www.ebi.ac.uk /Tools/clustalw2/index.html) to confirm that the inserts were identical apart from the base at rs4444903.

Restriction digestion of the TA constructs and destination vectors pGL3C, pGL3P and pGL3B was carried out in 10µl volumes with 1µl fast digest HindIII (Fermentas), 1µl fast digestion buffer, and 1 µg DNA for 10 min at 37°C. The reaction was heat inactivated at 80 °C for 10 min. De-phosphorylation of the linearized vector was carried out in the same reaction by addition of 1 U of Shrimp alkaline phosphatase in order to minimize vector recircularization. The fragments were gel purified using gel extraction kits (Qiagen). 100 ng of the linearized de-phosphorylated vector and 30ng of the insert (1:3 molar ratio) were subjected to a ligation reaction using 1 U T4 DNA ligase and 1x Ligation buffer in 10µl at room temperature for 1 hour. Transformation of X1-blue super-competent cells was carried out according to the manufacturer's protocol (Stratagene). The colonies containing the recombinant vector were selected by restriction enzyme treatment using HindIII, and were sequenced as above using GL2 primer to confirm the exact sequence and orientation. Endotoxin free minipreps (Promega) of the constructs containing the correct sequence were prepared (PGL3CA/CG, pGL3PA/PG, and pGL3BA/BG) (Figure 3-11).

Figure 3-11: Plasmid constructs used in the dual luciferase assay and in vitro transcription

a) 442 bp of the 5'UTR region of the EGF gene from A and G alleles were cloned into pGL3Basic (promoterless) vector to create BA and BG constructs. b) 390 bp representing the EGF promoter region was cloned into pGL3Basic vector to create EGFP construct. c) 832 bp representing 5'UTR region and EGF promoter was cloned in pGL3Basic vector to create BAL and BGL constructs. d) 442 bp of 5'UTR region was cloned in pGL3Promoter vector downstream to Simian virus promoter and upstream to luciferase gene to create PA and PG constructs. e) 442 bp of 5'UTR region was cloned in pGL3control vector downstream to Simian virus 40 promoter (SV40P) and upstream to luciferase gene and SV40enhancer (SV40E) to create CA and CG constructs. f) 5'UTR region of the EGF gene and luciferase gene were cloned from BA and BG constructs into pcDNA3.1 vector downstream to T7 promoter to create A-Luc and G-Luc respectively, to be linearized with Xba-1 before in vitro transcription. The Luc construct was created by cloning the luciferase gene only from pGL3Basic vector into pcDNA3.1 vector in the same way.

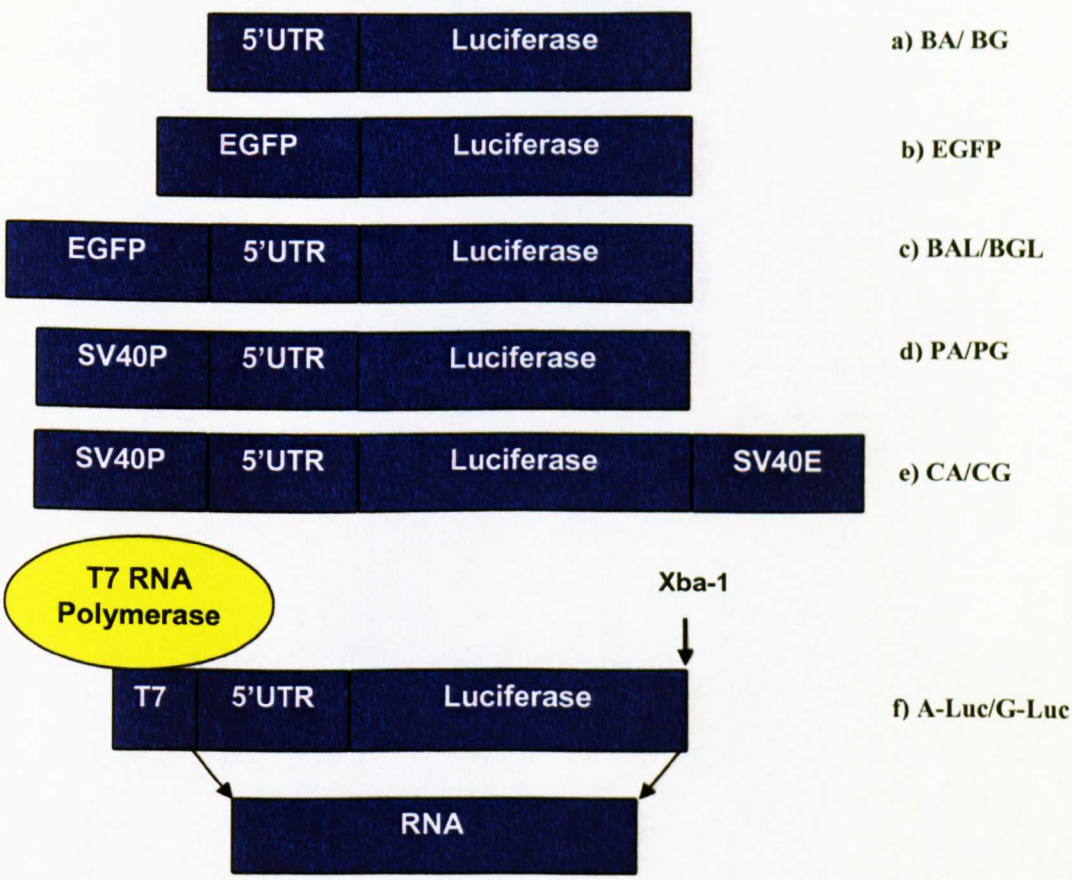


Table 3-7: Primers used for experiments in 3.5.1

HindIII site is underlined; AttB recombination sites are shown in bold.

Primer	Sequence
EGF-UTR-S	5'-TCAAAAAAGCTTACTGTTGGGAGAGGAATCGTA-3'
EGF-UTR-AS	5'-AGTGAGAAGCTTAATCTTGATGAGTTTGAACTTTCA-3'
EGF-PRO-S	5'GGGGACCACTTTGTACAAGAAAGCTGGGTGAACTAT CTTTACTATTGCTCATGT-3'
EGF-PRO-AS	5'GGGGACCACTTTGTACAAGAAAGCTGGGTTTCTCTT TTTGAAAAGTGGAAT-3'
EGF-PUTR-AS	5'GGGGACCACTTTGTACAAGAAAGCTGGGTAATCTTG ATGAGTTTGAACTTTC-3
M13 F primer	5'-TGTAAAACGACGGCCAGT-3'
M13R primer	5'-CAGGAAACAGCTATGAC-3'
GL2	5'-CTTTATGTTTTTGGCGTCTTCCA-3'
RV3	5'-CTAGCAAAATAGGCTGTCCC-3'
T7	5'-TAATACGACTCACTATAGG- 3'
BGH_rev	5'-TAGAAGGCACAGTCGAGG- 3'

3.5.3.2 Cloning of the EGF promoter and EGF promoter/5'UTR fragments

The Gateway cloning system takes advantage of the site-specific recombination reactions enabling the bacteriophage lambda to integrate and excise itself in and out of a bacterial chromosome. Gateway protocols rely essentially on the BP and LR clonase reactions described below (Karimi et al., 2007, Hartley et al., 2000).

390 bp representing the EGF promoter (-390>+1) were amplified from genomic DNA using the EGF-PRO-S and EGF-PRO-AS primer pair. 832 bp representing EGF promoter/ 5'UTR (-390>+442) were amplified using gDNA homozygous for the rs4444903 A and G alleles with the primer pair EGF-PRO-S and EGF-PUTR-AS. These primers are tagged at the 5'end by the AttB recombination sites (Table 3-7).

The PCR reaction was performed in 30µl volumes containing 200ng of gDNA, 1xTaq buffer containing 1.5 mM $MgCl_2$, 200 µM dNTPs, 0.5 µM of each primer, and 1 unit of Taq polymerase. The conditions were 94°C for 40 s, and 35 cycles of denaturing at 94°C for 30 s, primer annealing at 58°C for 1 min, and extension at 72°C for 1 min, with a final extension of 10 min at 72°C. PCR products were visualized by electrophoresis on 1% agarose gels stained with ethidium bromide.

1. BP reaction

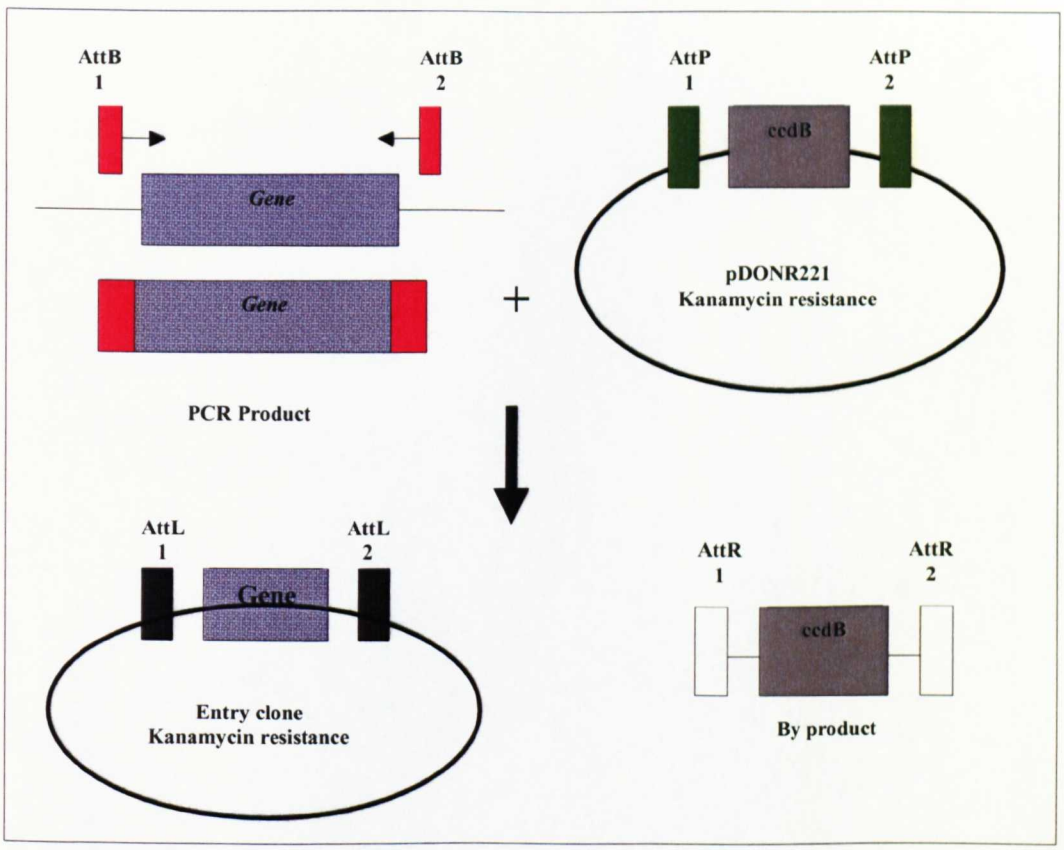
The principle of the BP reaction is illustrated in

Figure **3-12**. PCR products flagged by the AttB1 and AttB2 sites were cloned into the vector pDONR221 with AttP1 and AttP2 sites to create an entry clone. This recombination reaction was carried out using 150 fmol of the pDONR221 vector, 3µl PCR product, 2µl BP Clonase II enzyme mix containing the phage integrase and the integration host factor, made up to 8 µl with TE buffer pH8. The reaction was incubated at 25°C for 1 hour after which 1µl of proteinase K was added and incubated again for 10 min at 37°C.

DH5α cells were transformed using 3µl of this BP reaction and plated on LB plates with kanamycin antibiotic added (50µg/ml), and incubated overnight at 37°C.This creates entry clones carrying the gene of interest flanked by AttL1 and AttL2 recombination sites. The clones were tested by restriction digestion and sequencing using M13 forward and reverse primers (Table 3-7).

Figure 3-12: BP reaction

In the presence of clonase enzyme mix recombining AttB and AttP sites will produce AttL and AttR sites with the formation of entry clones. The ccdB gene is toxic for the bacteria so that nearly 90 % of the colonies created will be the correct colony, which minimizes colony screening.



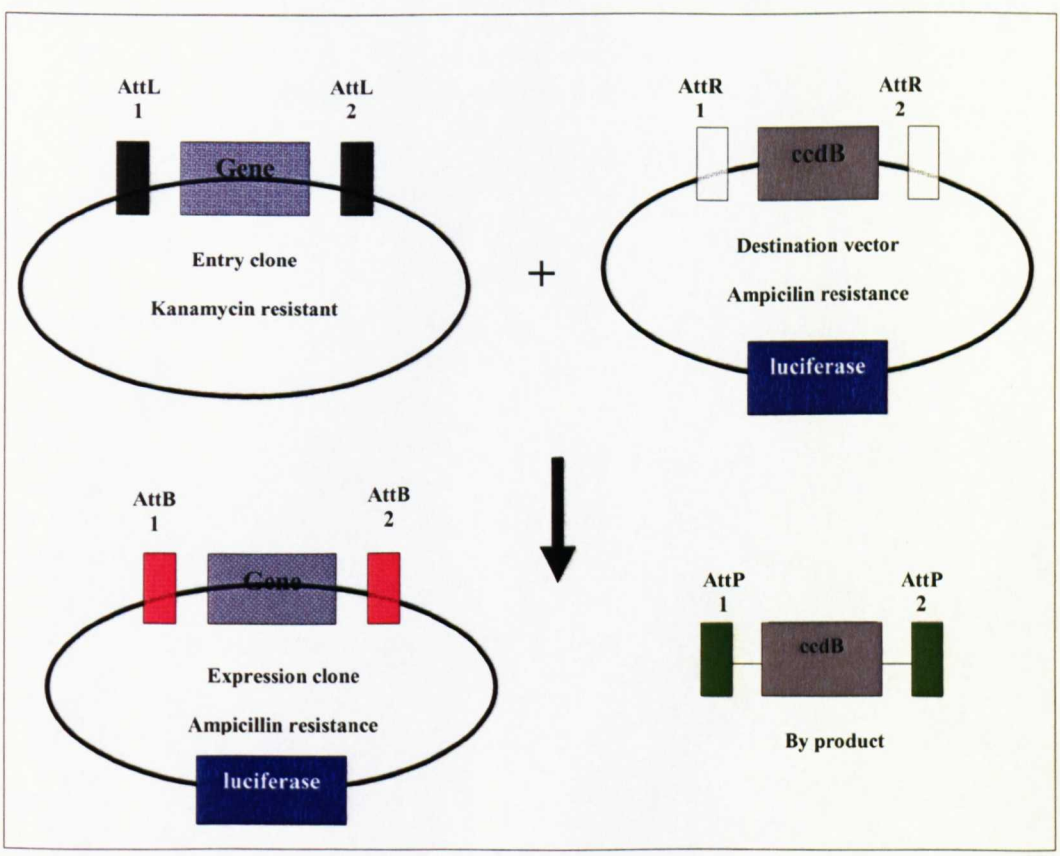
2. LR reaction

A second recombination reaction was carried out to create the final expression clones, using the pGL3B vector with AttR1 and AttR2 recombination sites as a destination vector and the entry clone from the BP reaction. The principle of the LR reaction is illustrated in Figure 3-13. 150 ng of each vector, 2µl of LR Clonase II enzyme mix containing integrase, integration host factor, and the phage excisionase, were mixed and made up to 8µl with TE buffer pH8. The

reaction was incubated at 25°C for 1 hour after which 1µl of proteinase K was added and incubated for 10 min at 37°C. This creates expression clones carrying the genes of interest flanked by AttB1 and AttB2 recombination sites. DH5α cells were transformed using 3µl of this LR reaction and plated on LB plates with Ampicillin antibiotic added (50µg/ml). Clones were tested by sequencing using GL2 and RV3 primers (Table 3-7). Endotoxin free minipreps for the constructs BAL, BGL and EGFP (Figure 3-11, B and C) were performed.

Figure 3-13: LR reaction

The recombination between AttL sites flanking the gene on the kanamycin resistant entry clones and AttR sites flanking the ccdB gene on the ampicillin resistant destination vector which carries the reporter gene creates an ampicillin resistant expression clone carrying both the gene of interest and the reporter gene.



3.5.3.3 Cell culture and transfection

1: Optimization for Jeg-3 cells transfection

Jeg-3 is a chorio-carcinoma cell line, which is highly proliferative and able to secrete hPL and hCG in culture media. These cells have many of the features of syncytiotrophoblast. Jeg-3 cells were maintained in Eagle's minimal essential medium (EMEM) with 10% FCS, 10 mM L-glutamine, penicillin (100U/ml), streptomycin (100 pg/ml), Amphotericin B 2.5 μg/ml and 1% sodium pyruvate at 37°C in 5% CO_2 in 75 cm^2 culture flasks. 2×10^5 Jeg-3 cells were plated in 12 well plates and cultured for 24 hours to attain 50-70% confluence before transfection. The cells were subcultured once they approached 80% confluence (usually after ~3 days of growth). Cells in a single 75 cm^2 flask were split equally into 3 new flasks as part of this process. As it was evident that expression of the reporter gene decreased as the cells were cultured for longer times, the passage number of the cells was limited to 5-6 passages, after which fresh aliquots were taken from the liquid nitrogen to establish consistency in the growth pattern and cellular characteristics over the experiments.

Two different transfection reagents, TFx20 and TFx50 (Promega), and various reagent to DNA ratios were tested to optimise the reaction. Serum free medium was used as recommended in a number of publications (Buckland et al., 2005, Hoogendoorn et al., 2003).

The transfection reagents were re-suspended in nuclease free water as specified in the manufacturer's protocol at least 24 hours prior to transfection. The reconstituted TFx reagents were stored in aliquots at -20^{0}C. 3, 4.5, and 6 ul of the reagent were used with 1 μg DNA to achieve 2:1, 3:1, and 4:1 charge ratios respectively.

Co-transfection was performed using 200ng of pGL3C and pGL3P vectors mixed with 20ng of pRL-SV40 (renilla expression vector) in 400μl serum free medium. TFx20 and TFx50 reagents were added according to the defined charge ratios. The reaction was vortexed briefly and incubated at room temperature for 15 min to allow the DNA to be incorporated into a complex with the lipid particles in the reagent so it can pass through the cell membrane. During this time the culture medium was removed from the cells and the cells were washed in 1X PBS. The TFx/DNA complex was then transferred to cells and incubated for 1 hour at 37°C after which 800μl of complete medium was added. All transfections were performed in triplicate in two independent experiments.

2: Transfection of HepG2 and Jeg-3 cells

HepG2 (hepatoma cell line) cells were maintained in culture as previously described (Section 3.4.2). 1× 10^6 HepG2 cells and 2× 10^5 Jeg-3 cells were plated in 12 well plates 24 hours before transfection to attain 50-70% confluence.

Co-transfection using 200ng of pGL3 experimental vectors (Figure 3-11, a-e) and 20ng of pRL-SV40 (renilla expression vector) was performed using TFx20

transfection reagent (Promega). Plasmid DNA and TFx20 transfection reagent in 3:1 charge ratio were incubated at room temptature in 400µl serum free medium for 15 min then transferred to cells and incubated for 1 hour at 37°C after which 800µl of complete medium was added. All transfections were performed in triplicate in 4-6 independent experiments.

Transfection with the pGL3C, pGL3P, and pGL3B empty vector was included in every experiment; a consistent fold increase from pGL3C to pGL3P vector over all over experiments was required as evidence of successful transfection.

3: Dual luciferase assay (DLA)

The Dual-Luciferase® Reporter Assay System (Promega) was used for reporter gene assays. All reagents were equilibrated to room temperature before use and prepared according to the manufacturer's protocol.

24 hours after transfection culture media were removed and the cells were washed 3 times with 1X PBS followed by addition of 200µl of 1X passive lysis buffer (PLB) and incubated on a shaker for 20 min at room temperature. 20µl of the cellular lysate was added to 50µl luciferase assay reagent II (LARII) in specific assay cuvettes suitable for the Turner Design Model TD-20/20 Luminometer. Firefly luciferase activity was measured then 50µl of 1X Stop and Glo reagent was added to the reaction for the measurement of renilla luciferase.

The results are expressed as relative luminescence units (RLU); the ratio of firefly luciferase/renilla luciferase is the normalized luciferase activity. The mean normalized luciferase activity of triplicate wells was used for statistical analysis.

3.5.3.4 In vitro transcription and translation

1: Cloning of constructs for in vitro transcription

BA, BG, pGL3B and pcDNA3.1 were subjected to restriction digestion using fast digestion Kpn I and Xba I enzymes following standard techniques. The DNA constructs representing the A and G alleles of the 5'UTR region of EGF with the luciferase gene, or the luciferase alone, were cloned into the pcDNA3.1 vector downstream to the T7 promoter using T4 DNA ligase as previously described (Section 3.4.3) to generate A-Luc, G-Luc and Luc constructs (Figure 3-11, f).

After confirmation of the constructs by restriction digestion using Kpn-I and Xba I and sequencing using T7 and BGH-rev primers (Table 3-7), midipreps (Qiagen) were carried out to generate large amounts of DNA. 10 µg of DNA were linearized using 5 U of fast digest Xba I and 1x fast digestion buffer in 50µl reaction volumes. Linearization was confirmed by running 2µl of the restriction digestion reaction on 1% agarose gel alongside the corresponding uncut construct. Purification of the linearized DNA was carried out by extraction with TE-saturated (pH 8.0) phenol:chloroform: isoamyl alcohol (25:24:1) and ethanol precipitation according to the manufacture's protocol provided with the in vitro transcription kit. The samples were re-suspended in 50µl Rnase-Free water and stored at –20°C.

2: In vitro transcription (IVT)

RiboMAX™ Large Scale RNA Production System T7 Cat # P1300 (Promega) was used for in vitro transcription. The IVT reaction was prepared at room temperature by adding 10µl of 5x transcription buffer, 15 ul of 25 mM rNTPs, 4 µg linearized constructs, 10µl T7 enzyme, and made up to 50µl with Rnase free water in a 0.5 ml microcentrifuge tube. The reaction was gently mixed and incubated at 37^0C for 3 hours. A positive control reaction using the commercially supplied linear DNA was included in the experiment.

RQ1-RNase free DNase at a concentration of 1U/µg of template DNA was used for removal of the unused DNA template following transcription. The reaction was incubated for 15 min at 37°C.

To confirm a single homogenous RNA product of the desired size, 2µl of IVT reaction (RNA) were added to 18µl of RNA sample buffer (10.0 ml deionized formamide, 3.5 ml 37% formaldehyde and 2.0 ml 5x MOPS buffer). 5µl of RNA loading buffer (50% glycerol, 1mM EDTA, and 0.4% bromophenol blue) were added and the sample was heated for 5–10 min at 70°C. The sample was then loaded onto a 1% agarose gel to which 0.5 µg/ml ethidium bromide was added. Electrophoresis was carried out at 80 V for 30 min. The control reaction that produces a transcript of approximately 1,800 bp, the same size as the Luc control, was used for size comparison.

3: RNA purification after in vitro transcription:

RNA was extracted from 30µl of the IVT reaction with 1 volume of citrate-saturated phenol (pH 4.7): chloroform: isoamyl alcohol (125:24:1) and ethanol precipitation according to the manufacturer's protocol. RNA was then suspended in 30µl RNase free water. RNA was purified from the remaining 20µl of the IVT reaction using RNA clean spin columns (RNeasy Mini Kit- Qiagen, Cat# 74104) according to a standard protocol. These two methods of RNA purification were compared for efficiency of removal of contaminants that can interfere with the subsequent translation experiment. RNA concentration was measured on the Nanodrop spectrophotometer.

4: In vitro translation

The template mRNA was denatured by heating at 65°C for 3 min and immediate cooling in an ice-water bath. The reaction was then assembled in a 0.5 ml polypropylene microcentrifuge tube by adding 35µl Rabbit Reticulocyte Lysate, 0.5µl of 1mM Amino Acid Mixture Minus Leucine, 0.5µl of 1mM Amino Acid Mixture Minus Methionine, 1µl of RNasin® Ribonuclease Inhibitor (40U/µl), and 2µg of the RNA constructs (A-Luc,G-Luc, and Luc), and made up to 50µl with RNase free water. The reaction was then incubated at 30°C for 90 min. One reaction using Luciferase Control RNA (1µg/µl) supplied with the kit was included as a positive control. A reaction without any added RNA was included as a negative control to detect any luminescence from the reaction components.

Luciferase output was measured using a luminometer by adding 2.5µl of the 50µl translation reaction to the 50µl of LARII. Each sample was measured in triplicate. The luciferase output from the Luc construct was above the upper limit of detection of the luminometer, so 1/5, 1/10, and 1/20 dilutions were prepared from 2µl of the translation reaction with the appropriate amount of nuclease free water prior to determination of luciferase activity.

3.5.3.5 Electrophoretic mobility shift assay, EMSA

HepG2 nuclear extract was prepared using NucBuster™ Protein Extraction Kit (Novagen) according to the instructions provided by the manufacturer. Double stranded DNA oligonucleotides were prepared by mixing 1µl of 100 pmol /ul of the forward and reverse probes (1:1 molar ratio) with 98µl of annealing buffer (10 mM Tris, 1 mM EDTA, 50 mM NaCl , pH 8.0) to give a final concentration of 1 pmol/µl of double stranded oligonucleotides. Annealing of oligonucleotides was carried out by incubating at 95°C for 5 min then gradually reducing the temperature to room temperature. A working concentration of 10fmol/µl was prepared immediately before performing the electrophoretic mobility shift assay. The oligonucleotide probes used were:

5'UTRA: 5'-AATCCAAGGGTTGTAGCTGGAACTTTCCAT-3'

5'UTRG: 5'-CAATCCAAGGGTTGTGGCTGGAACTTTCCAT-3'

Sp1: 5'-GTCCAAGCTCCCCGCCCCTCCCCAGC-3'.

The probes were biotinylated at the 5' end. Unbiotinylated specific probes and a non specific scrambled probe were also used in these experiments. Specificity protein 1 (Sp1) antibody was supplied by Milipore (Anti-Sp1, Cat # 07-645).

EMSAs were set up according to the protocol provided by the manufacturer (Thermo Scientific, LightShift™ Chemiluminescent EMSA Kit, Product No. 20148), shown in Table **3-8.** Briefly, HepG2 nuclear extract was incubated with 20 fmol biotin-labelled double stranded DNA probe on ice for 20 min in a 20µl reaction volume, with or without inclusion of competing specific or non-specific oligonucleotides. 10µl of the reaction was then loaded onto a 6% polyacrylamide gel and subjected to electrophoresis in 0.5% TBE (Tris-borate EDTA) buffer. For super-shift assays, 1µg of Sp1 antibody was incubated for 10 min on ice with reactions containing nuclear extract before adding the double stranded DNA probe followed by 20 min incubation on ice prior to electrophoresis. The binding reactions were transferred to a nylon membrane using a semi-dry blotting system and cross-linked by exposure to UV light for 15 min. Detection of the biotin-labelled DNA was carried out using a chemiluminescent method. The membrane was incubated with Streptavidin-Horseradish peroxidise solution, followed by addition of stable peroxide solution according the manufacturer's protocol. The bands then were viewed on X-ray film.

Table 3-8: Electrophoretic mobility shift assay for 5'UTRA, 5'UTRG of the EGF gene and Sp1 probes

Reactions are labelled 1-6 as follows: 1=10fmol/μL biotin end-labelled test probe; 2=10fmol/μL biotin end-labelled test probe and HepG2 nuclear extract; 3=10fmol/μL biotin end-labelled test probe, HepG2 nuclear extract and 100 molar excess unlabeled non specific probe; 4=10fmol/μL biotin end-labelled test probe, HepG2 nuclear extract and 100 molar excess unlabeled specific probe; 5=10fmol/μL biotin end-labelled test probe, HepG2 nuclear extract and Sp1 antibody; 6= HepG 2 nuclear extract alone.

Component	Final concentration	Reaction*					
		1	2	3	4	5	6
Ultrapure Water		15 μL	13 μL	11 μL	11 μL	12 μL	15 μL
10X Binding Buffer (20148A)	1X	2μL	2μL	2μL	2μL	2μL	2μL
1μg/μL Poly (dI•dC) (20148E)	50 ng/μL	1μL	1μL	1μL	1μL	1μL	1μL
HepG2 cells nuclear extract	Optimized	-	2μL	2μL	2μL	2μL	2μL
1pmol/μL unlabeled non specific Target DNA	2 pmol	-	-	2 μL	-	-	-
1pmol/μL unlabeled specific Target DNA		-	-		2 μL	-	-
Sp1antibody 1μg/1μL		-	-	-	-	1μL	-
10fmol/μL biotin end-labelled test DNA	20 fmol	2μL	2μL	2μL	2μL	2μL	-
Total Volume	----	20μL					

3.5.3.6 Bioinformatics

Mfold program was used for prediction of the EGF 5'UTR secondary structure (http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi) (Zuker, 2003) to determine which allele is expected to translate more efficiently. The prediction was performed for the full length and first 120 bases only of the 5'UTR region. Prediction of transcription factor binding using the TF search program (http://www.cbrc.jp/research/db/TFSEARCH.html) was also carried out. The

input was 30 bp of the EGF 5' UTR with the c.61 A>G SNP in the middle. The default parameter of the program was applied, where the lower limit of the score was set to 70 points, implying the degree of correlation between the input sequence and the consensus sequence for the transcription factor binding included in the program.

3.5.3.7 Data analysis

In the dual luciferase assay, the mean of the relative luciferase activity (firefly/renilla) for each construct was used. Data were rejected whenever the coefficient of variation (CV) between the triplicates exceeded 20%. An average fold change the range of 95 and 14 for pGL3C/pGL3P was required all the experiments for Jeg-3 and HepG2 cells respectively. Each construct was then corrected to its corresponding empty vector to obtain the activity related to the construct itself. The empty pGL3B vector showed very low activity in Jeg-3 cells, so pGL3B constructs were corrected to the empty pGL3P vector. The luciferase outputs for each allele were compared. To test whether 5'UTR constructs in pGL3B conferred any promoter activity, both were compared with the activity of the EGF promoter in the EGFP construct (Hoogendoorn et al., 2003). Comparisons between construct activities were carried out using Student's t test for data with a normal distribution, or Mann Whitney U test for data with a non-normal distribution.

3.5.4 Results

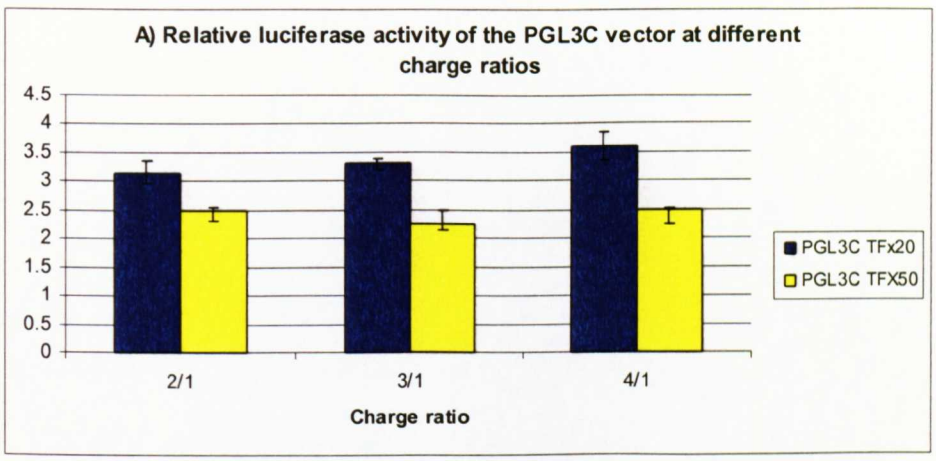
3.5.4.1 Optimization of Jeg-3 cells transfection conditions

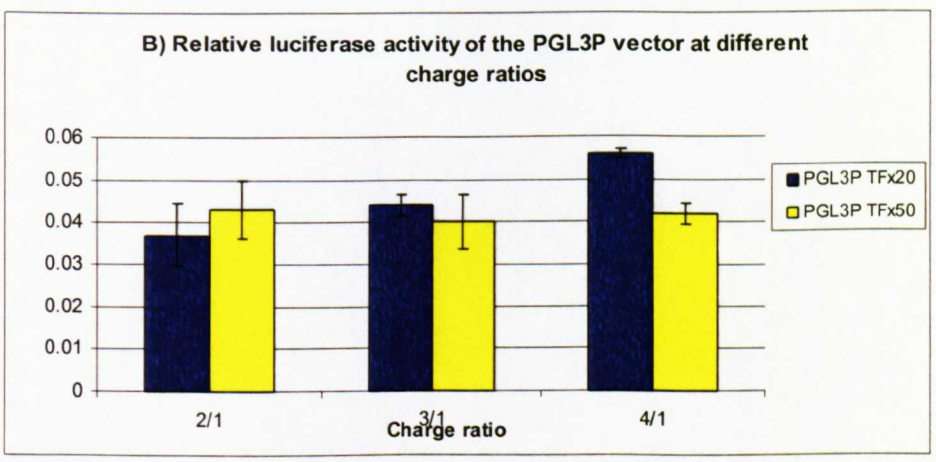
Use of TFx20 gave higher relative luciferase activity compared to TFx50 at all charge ratios with both pGL3C and pGL3P vectors (Figure 3-14). So TFx20 was used for Jeg-3 transfection experiments. The fold change in pGL3C/PGL3P activity was examined to determine the optimum charge ratio. TFx20 at a charge ratio of 3:1 showed the highest fold change averaged over the two experiments. Taking into account that the higher the fold change the better the experimental conditions to discriminate between different constructs, a charge ratio of 3:1 was adopted for the subsequent experiments (Figure 3-14).

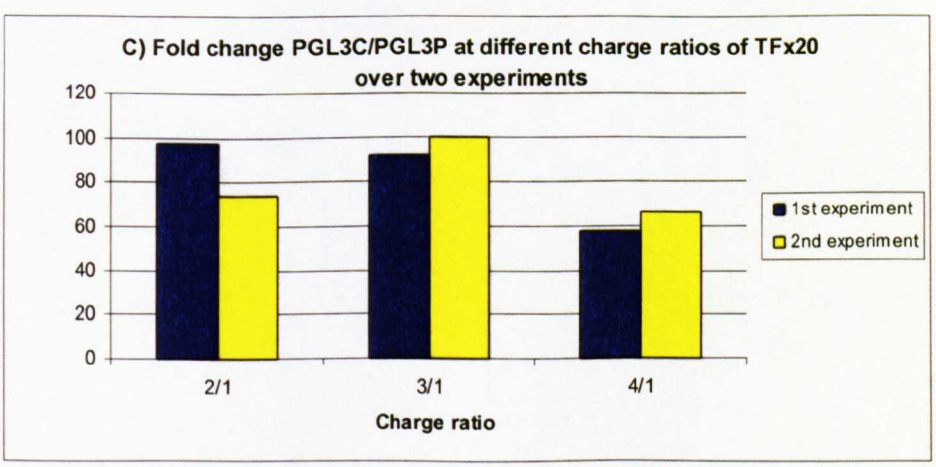
According to the Promega notes a pGL3C/PGL3P fold change of between 6 and 136 is acceptable and depends on the cell line under investigation. Based on the results of optimisation experiments, a pGL3C/PGL3P fold change near to 95 was considered acceptable for subsequent experiments. The coefficient of variation (CV%) of relative luciferase activity under these experimental conditions was 11% for pGL3C and 17% for pGL3P. This is comparable to previous studies (Hoogendoorn et al., 2003, Buckland et al., 2005).

Figure 3-14: Optimization of the Jeg-3 cells transfection

Transfection of $2x10^5$ cells with 200ng of pGL3C (A), and pGL3P (B) along with 20ng of pRL renilla at different charge ratio using TFx20 and TFx50 reagents. Transfection was run in triplicate, data represent mean ± SEM for 2 experiments. pGL3C/pGL3P fold change (C).







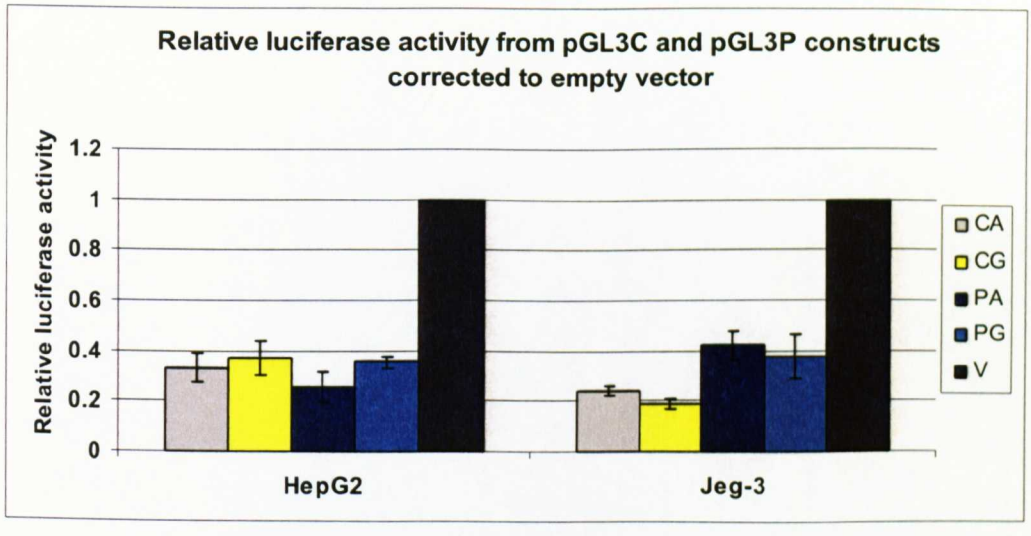
The transfection conditions for HepG2 cells had been optimized previously in our laboratory. TFx20 at ratio 3:1 was used. The average fold change of pGL3C/PGL3P was 14. A 20% coefficient of variation was accepted as maximum variations between the replicates in subsequent experiments.

3.5.4.2 Effect of EGF 5'UTR on reporter gene expression

The pGL3C and pGL3P vectors have high basal transcription, so that any variation in reporter gene expression is expected to be due to differences in the effects of inserts on translation rather than transcription. Interestingly, the relative luciferase activity of both A and G alleles of the 5' UTR of the EGF gene constructs was between 20 to 40% of the pGL3C and pGL3P empty vectors in both cell lines, implying that the 5'UTR down-regulates gene translation. Constructs containing the 442 bp 5'UTR of the EGF gene upstream to the luciferase gene in pGL3C (CA to CG) and pGL3P (PA to PG) showed that although the G allele tended to be more active than the A allele in HepG2 cells and less active in Jeg-3 cells, the differences were not statistically significant (Figure 3-15 and Table 3-9), with no evidence that this variant affects the efficiency of translation.

Figure 3-15: Relative luciferase activity for the EGF 5'UTR A and G alleles in pGL3C and pGL3P

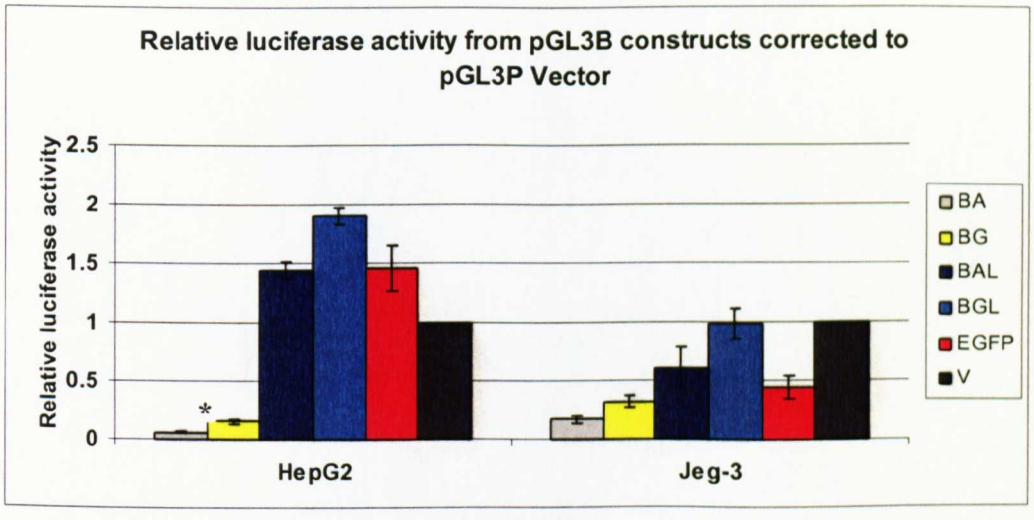
Data are presented as mean ± SEM of 5 independent experiments from the 2 cell lines under investigation. CA and CG, PA and PG: 442 bp EGF 5'UTR from the A and G alleles inserted into pGL3C and pGL3P vector respectively. V: empty pGL3C or pGL3P vector.



In pGL3B constructs, in which the 442 bp 5'UTR of the EGF gene was inserted upstream to the luciferase gene in the promoter-less pGL3B vector, there was higher expression from the G allele compared to the A allele in HepG2 (P= 0.004) but not in Jeg-3 cells (P= 0.08), (Figure 3-16). Inclusion of 390bp of the EGF promoter together with the 442 bp 5'UTR in the BAL and BGL constructs, inserted into the pGL3B vector, showed a non-significant difference in the activity of the G allele compared to A allele in HepG2 cells (P=0.1) and Jeg-3 cells (P=0.09) (Figure 3-16 and Table 3-9).

Figure 3-16: Relative luciferase activity from EGF 5'UTR A and G alleles in pGL3B constructs

Data are presented as mean ± SEM of 5 independent experiments from the 2 cell lines under investigation. BA and BG: 442 bp EGF 5'UTR from the A and G alleles inserted into pGL3B vector. BAL and BGL: 442 bp EGF 5'UTR from the A and G alleles and the 390 bp of the EGF promoter inserted in pGL3B. EGFP: 390 bp of the EGF promoter inserted in pGL3B vector and V= empty vector.* BA compared to BG in HepG2 cells (P = 0.004).



In HepG2 cells, the mean CG/CA, PG/PA, BG/BA, and BGL/ BAL fold change in the luciferase activities were 1.10, 1.55, 2.59 and 1.38, respectively. CG/CA, PG/PA, BG/BA and BAL/BGL fold change in the luciferase activities in Jeg-3 cells were 0.80, 0.88, 1.85 and 1.65 respectively, (Table 3-9 and Figure 3-17) . The increment in the G/A fold change in the pGL3B compared with the pGL3C and pGL3P constructs indicates that the variation between A and G alleles is mostly transcriptional, and is masked when basal transcription is high.

Figure 3-17: Fold change in the activity of G allele/A allele of the 5'UTR of the EGF gene in different pGL3 constructs

Data are presented as mean ± SEM of 5 independent experiments from the 2 cell lines under investigation. CA and CG, PA and PG, BA and BG: 442 bp EGF 5'UTR from the A and G alleles inserted in pGL3C, pGL3P and pGL3B vector respectively. BAL and BGL: 442 bp EGF 5'UTR from the A and G alleles and the 390 bp of the EGF promoter inserted in pGL3B.

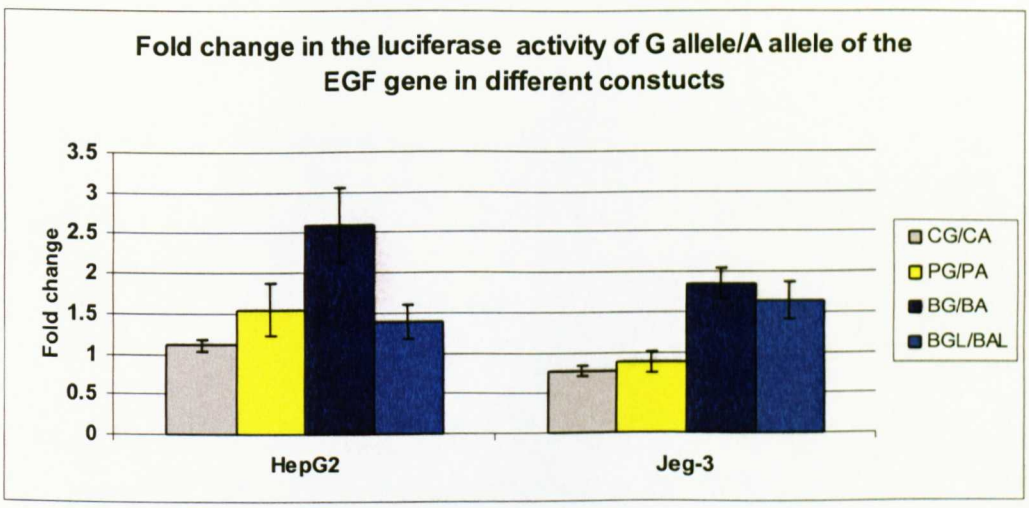


Table 3-9: Relative luciferase activity for A and G alleles of the EGF 5'UTR in different pGL3 vectors in HepG2 and Jeg-3 cells.

The data are presented as mean (SEM) and Student's t test P value, followed by fold change in the activity of G/A alleles, mean (SEM). CA and CG: 442 bp EGF 5'UTR from the A and G alleles inserted in pGL3C vector; PA and PG: 442 bp EGF 5'UTR from the A and G alleles inserted in pGL3P vector; BA and BG: 442 bp EGF 5'UTR from the A and G alleles inserted in pGL3B vector; BAL and BGL: 442 bp EGF 5'UTR from the A and G alleles and the 390 bp of the EGF promoter inserted in pGL3B.

	Relative luciferase activity	
	HepG2	Jeg-3
CA	0.33 (0.06)	0.24 (0.02)
CG	0.37 (0.07)	0.19 (0.02)
P value	0.69	0.11
Fold change CG/CA	**1.1 (0.07)**	**0.80 (0.07)**
PA	0.26 (0.05)	0.42 (0.06)
PG	0.36 (0.07)	0.38 (0.08)
P value	0.22	0.63
Fold change PG/PA	**1.55 (0.33)**	**0.88 (0.13)**
BA	0.06 (0.01)	0.18 (0.06)
BG	0.14 (0.05)	0.33 (0.07)
P value	0.004	0.08
Fold change BG/BA	**2.59 (0.46)**	**1.85 (0.19)**
BAL	1.44 (0.17)	0.61 (0.17)
BGL	1.90 (0.14)	0.98 (0.12)
P value	0.1	0.09
Fold change BGL/BAL	**1.38 (0.20)**	**1.65 (0.23)**

When comparing the transcriptional activity of BA and BG constructs (442 bp 5'UTR region) to EGFP (390 bp of the EGF promoter) inserted in pGL3B (promoter-less vector), both the A and G alleles had extremely low activity compared to the EGF promoter in HepG2 cells ($P< 0.001$ comparing either BA or BG to the EGFP). Interestingly, in Jeg-3 cells the A allele of the 5'UTR was considerably less active than the EGFP ($P= 0.02$), but the G allele had comparable activity to the EGFP ($P= 0.25$), (Figure 3-16 and Table 3-10).

Comparing BAL and BGL (constructs containing both 5'UTR and EGF promoter) to the EGFP (construct containing the EGF promoter alone) showed no difference in the activity of BAL and BGL compared to EGFP (P= 0.95, and 0.08 respectively) in HepG2 cells. In Jeg-3 cells, BAL showed no difference in activity while BGL had significantly higher activity compared to the EGFP construct (P= 0.22 and 0.02 respectively), (Figure 3-16 and Table 3-10)

In HepG2, the mean fold change in the luciferase activity of BA, BG, BAL, and BGL over EGFP were 0.05, 0.1 0.99, and 1.34, indicating that the 5'UTR A allele confers no transcriptional activation to this DNA fragment additional to the EGFP, while the G allele activates transcription. In Jeg-3 cells the fold change of BA, BG, BAL and BGL over the EGFP were 0.39, 0.74, 1.4 and 2.37 respectively, indicating that in Jeg-3 cell both A and G alleles have transcriptional activity but the G allele is transcriptionally more active Figure 3-18 and Table 3-10) .

Figure 3-18: Fold change in the activity of EGF 5'UTR A and G alleles with and without the EGF promoter relative to the EGF promoter alone

Data are presented as mean ± SEM of 5 independent experiments from the 2 cell lines under investigation. BA and BG: 442 bp EGF 5'UTR from the A and G alleles inserted into pGL3B vector; BAL and BGL: 442 bp EGF 5'UTR from the A and G alleles and 390 bp of the EGF promoter inserted into pGL3B; EGFP: 390 bp of the EGF promoter inserted into pGL3B vector.

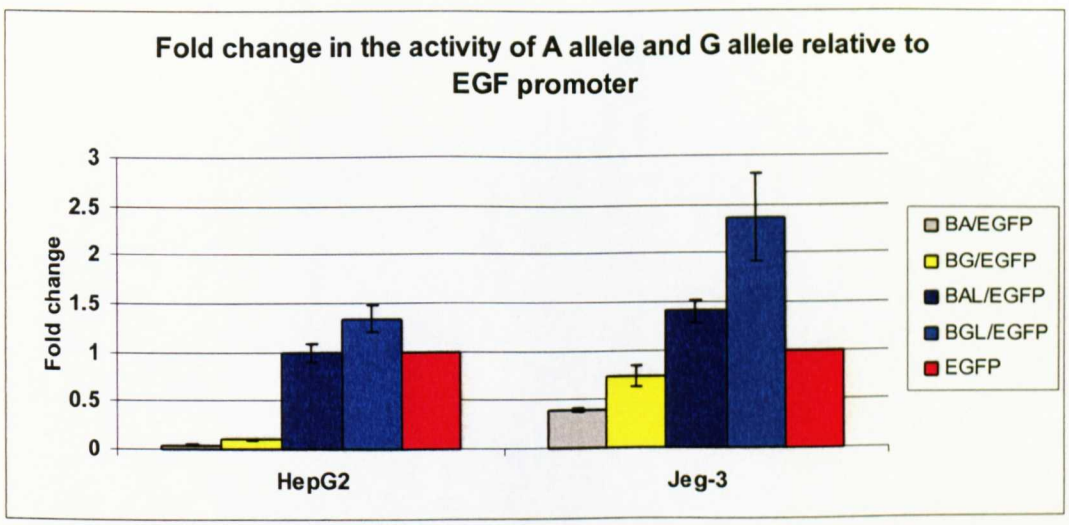


Table 3-10: The activity of EGF 5'UTR A and G alleles with and without the EGF promoter relative to the EGF promoter alone

The data are presented as mean (SEM) and Student's t test P value, followed by mean (SEM) fold change in the activity of A and G alleles in relation to the EGFP, BA and BG: 442 bp EGF 5'UTR from the A and G alleles inserted in pGL3B vector; BAL and BGL: 442 bp EGF 5'UTR from the A and G alleles and the 390 bp of the EGF promoter inserted in pGL3B.; EGFP: 390 bp of the EGF promoter inserted in pGL3B vector.

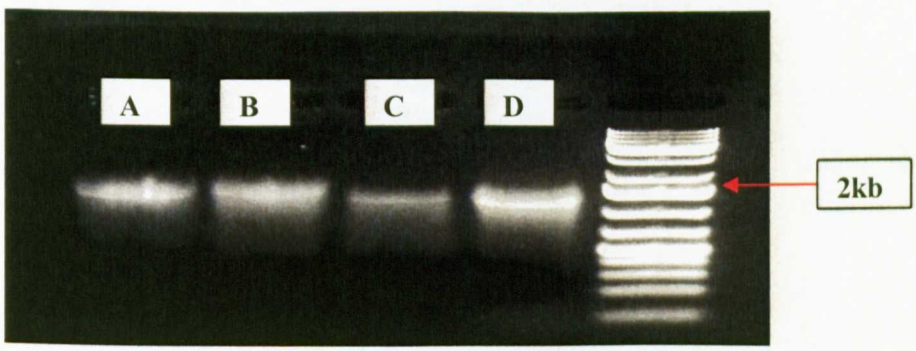
	Relative luciferase activity	
	HepG2	**Jeg-3**
EGFP	1.46 (0.16)	0.44 (0.07)
BA	0.06 (0.01)	0.18 (0.06)
P value	**[<0.001]**	**[0.02]**
Fold change BA/EGFP	0.05 (0.01)	0.39 (0.02)
BG	0.14 (0.05)	0.33 (0.07)
P value	**[< 0.001]**	**[0.25]**
Fold change BG/EGFP	0.10 (0.01)	0.74 (0.11)
BAL	1.44 (0.19)	0.62 (0.098)
P value	**[0.95]**	**[0.22]**
Fold change BAL/EGFP	0.99 (0.09)	1.40 (0.12)
BGL	1.90 (0.14)	0.98 (0.173)
P value	**[0.08]**	**[0.02]**
Fold change BGL/EGFP	1.34 (0.13)	2.37 (0.45)

3.5.4.3 In vitro transcription and translation

In vito transcription of the A-Luc and G-Luc constructs generated a transcript of 2242 bp, while transcription of Luc construct and the RNA control supplied by the Kit genetrated a 1800 bp transcript (Figure 3-19).

Figure 3-19: In vitro transcription reaction on 1% agarose gel

A and B are RNA products of 2242 bp from A-Luc and G-Luc constructs. C and D are 1800 bp RNA products from the Luc construct and the commercial RNA control. RNA loaded directly from the transcription reaction prior to purification.



In vitro translation of equal amounts of luciferase RNA under the control of the A and G alleles of the 5' UTR region of EGF gene showed that both alleles are translated with the same efficiency independent of the cellular environment, indicating that this sequence variation from A>G has no impact on translation of the down stream gene (Figure 3-20). The RNA was purified after transcription by phenol:chloroform:isoamyl alcohol (125:24:1) followed by ethanol precipitation (PPT) and by RNA clean spin column (Spin). The luciferase output was higher with the second method of purification, but there was no difference between the translational activity of the A and G alleles (P=0.86) and (P=0.76) for RNA purified by PPT or by Spin column methods respectively.

Both the A and G alleles lowered luciferase output by nearly 70% when introduced upstream of the luciferase gene compared to the output from the luciferase gene alone (P= 0.13 and 0.09 comparing A-Luc and G-Luc to Luc constructs prepared by PPT and SPIN methods respectively) (Figure 3-21). This experiment confirmed the findings of the reporter gene assay in pGL3C or pGL3P vectors that the 5'UTR region reduces the translation efficiency of the downstream gene.

Figure 3-20: Luciferase output from A and G alleles of the 5'UTR of the EGF gene after in vitro transcription and translation

Data are presented as mean±SEM of the relative luminescence units (RLU) from 4 independent experiments. A-Luc and G-Luc are the mRNA comprise of 5' UTR from the A allele and G alleles and the luciferase gene. PPT: RNA purified by precipitation, SPIN: RNA purified by spin column.

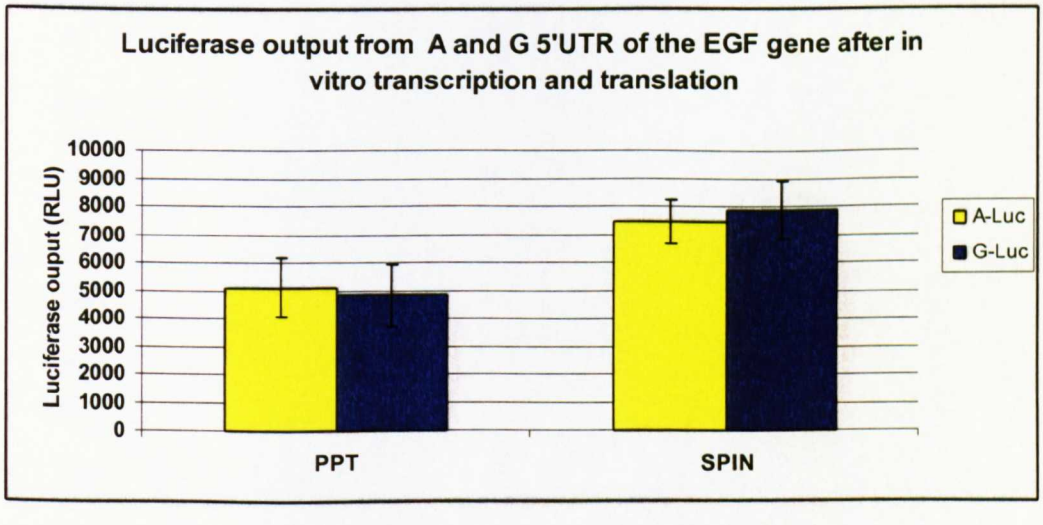
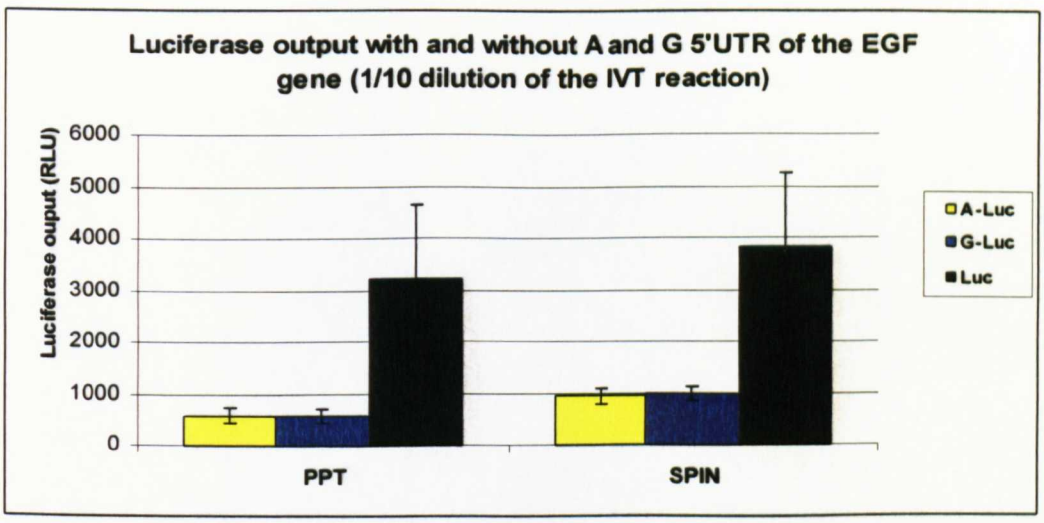


Figure 3-21: Luciferase output showing the effect of EGF 5'UTR A and G alleles on the translation of the downstream gene

Data are presented as mean ±SEM of the relative luminescence units (RLU) of a 1/10 dilution of the in vitro translation reaction from 4 independent experiments. A-Luc and G-Luc comprise the 5' UTR from the A or G alleles and the luciferase gene; Luc: luciferase gene; PPT: RNA purified by precipitation; SPIN: RNA purified by spin column.



Bioinformatic examination of the EGF-5'UTR A and G alleles found them to be identical in their mRNA secondary structure, other than the loop carrying A is somewhat less stable compared to G because of internal complementarity (shown in Figure 2-22 with the G allele boxed in red). The 5'UTR region of the EGF gene is highly structured with multiple stem loops that may interfere with the efficiency of translation (Figure 3-23). It also has 3 uORF upstream of the physiological start site.

Figure 3-22: mRNA secondary structure of the first 120bp of the EGF 5'UTR.

A (upper panel) and G (lower panel) alleles with SNP boxed in red. This mRNA secondary structure predicted using Mfold software.

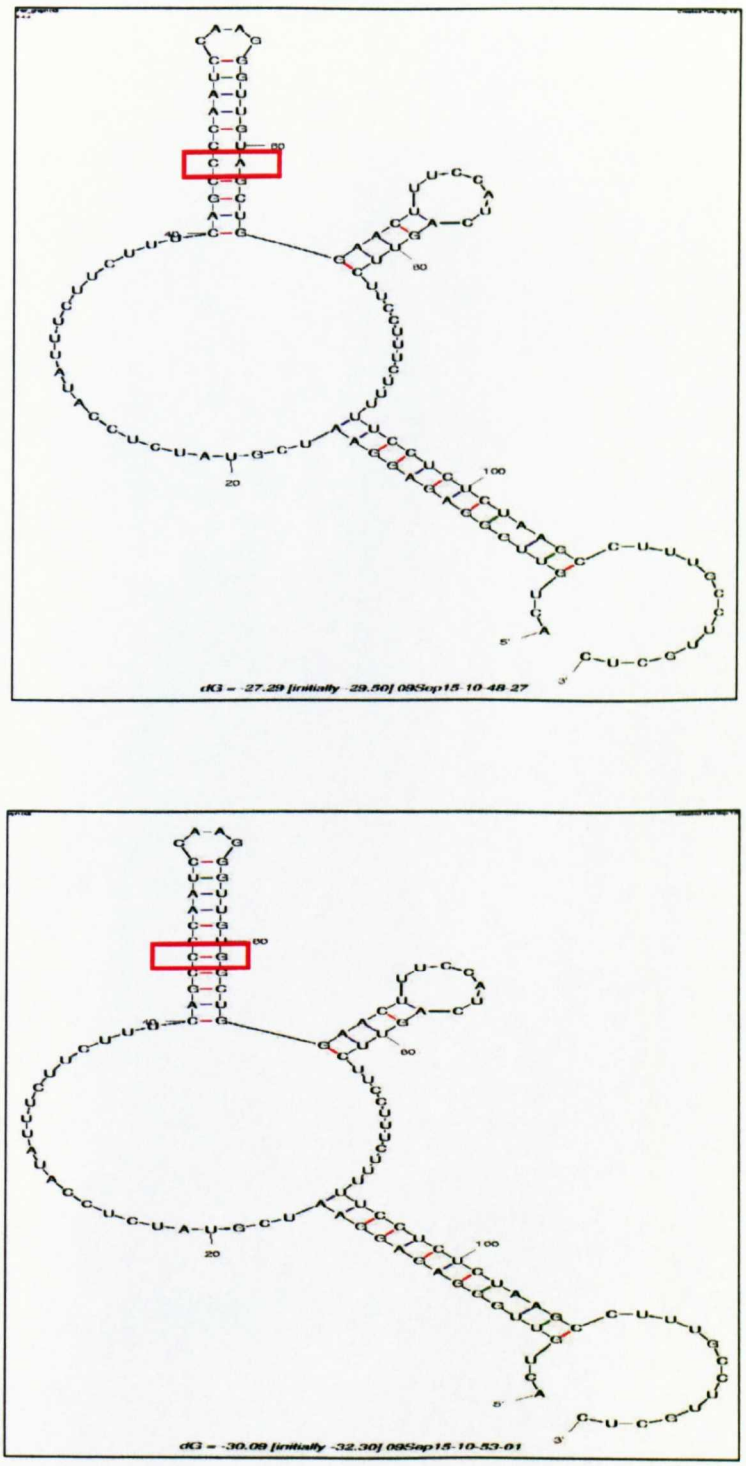
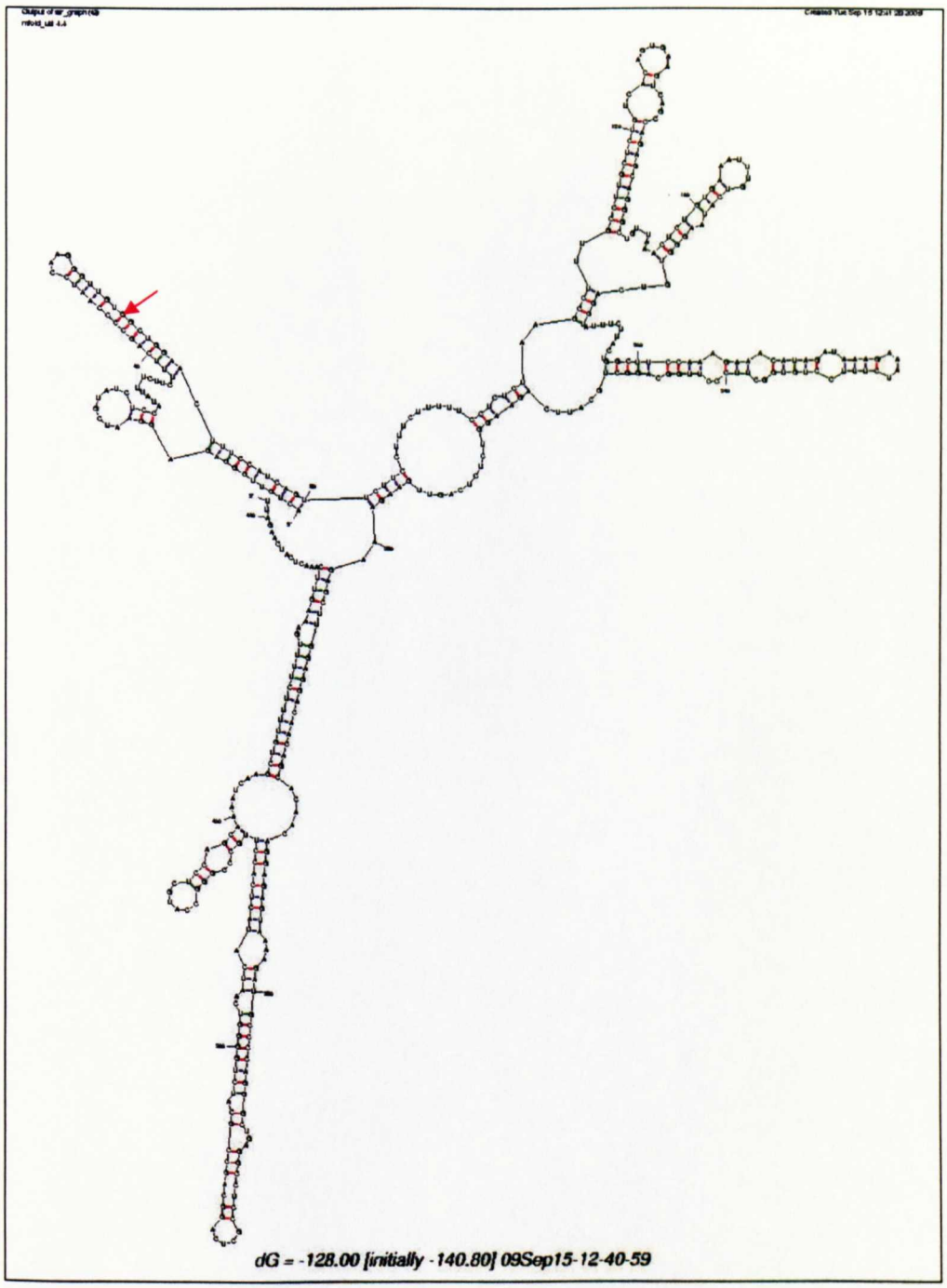


Figure 3-23: mRNA secondary structure of the full length EGF-5'UTR

The red arrow refer to the SNP, both A and G alleles were identical when the entire 442 bp 5'UTR was tested. This mRNA secondary structure predicted using Mfold software.



3.5.4.4 Electrophoretic mobility shift assay

Transcription factor binding site prediction showed that AML-1a, Sp1 and P300 transcription factors can bind exclusively to the G allele. Three other factors, Th/E4, Elk-1 and STATx, are predicted to bind to both A and G alleles (Figure 3-24).

Figure 3-24: Prediction of the transcription factors binding to 5'UTR of the EGF.

Output from the TF search program. The input was 30 bp of the EGF 5' UTR with the rs4444903 (A>G) SNP in the middle.

Transcription factor binding to G allele

```
G CAATCCAAGG GTTGTGGCTG GAACTTTCCA T
             ----->
          ------------------>
               ---------->
               ---------------->
       --------------->
               <---------
```

entry		score
M00271	AML-1a	83.7
M00222	Th1/E4	76.5
M00008	Sp1	75.3
M00025	Elk-1	73.0
M00033	p300	72.3
M00223	STATx	72.1

Transcription factor binding to A allele

```
A CAATCCAAGG GTTGTAGCTG GAACTTTCCA T
             ---------------->
          ------------------>
               <---------
```

entry		score
M00025	Elk-1	74.3
M00222	Th1/E4	74.1
M00223	STATx	71.2

Binding to Sp1 in HepG2 nuclear extract was investigated by EMSA assay. The results are shown in Figures 3-27 and 3-28. Both the G and the A alleles showed a band that is comparable to the band created when the Sp1 consensus sequence was used. The binding was specific for both G and A alleles, such that the use of 100 molar excess of unlabelled non specific sequence did not abolish the band (Figure 3-25), while the use of 100 molar excess of unlabelled specific competitor abrogates the binding (Figure 3-26). In the case of the A allele the band was removed by adding Sp1 antibody to the binding reactions, an effect which was also observed with Sp1 consensus sequence, but not observed with the G allele (Figure 3-26).

Figure 3-25: Electrophoretic mobility shift assays for 5'UTR of the EGF gene, and Sp1 site

The arrows indicate protein–DNA interactions. Band A was observed with the three probes used. This band most probably corresponds to Sp1. Band B was observed with both 5'UTRA and G, while band C was observed only with 5'UTRA. All of these bands were specific as they persisted in the presence of 100 molar excess of the unlabeled non specific probe. Non-specific bands are observed at the top of the gel due to endogenous biotinylated materials in the nuclear extract, which are also observed when nuclear extract only is loaded onto the gel (extreme right lane).

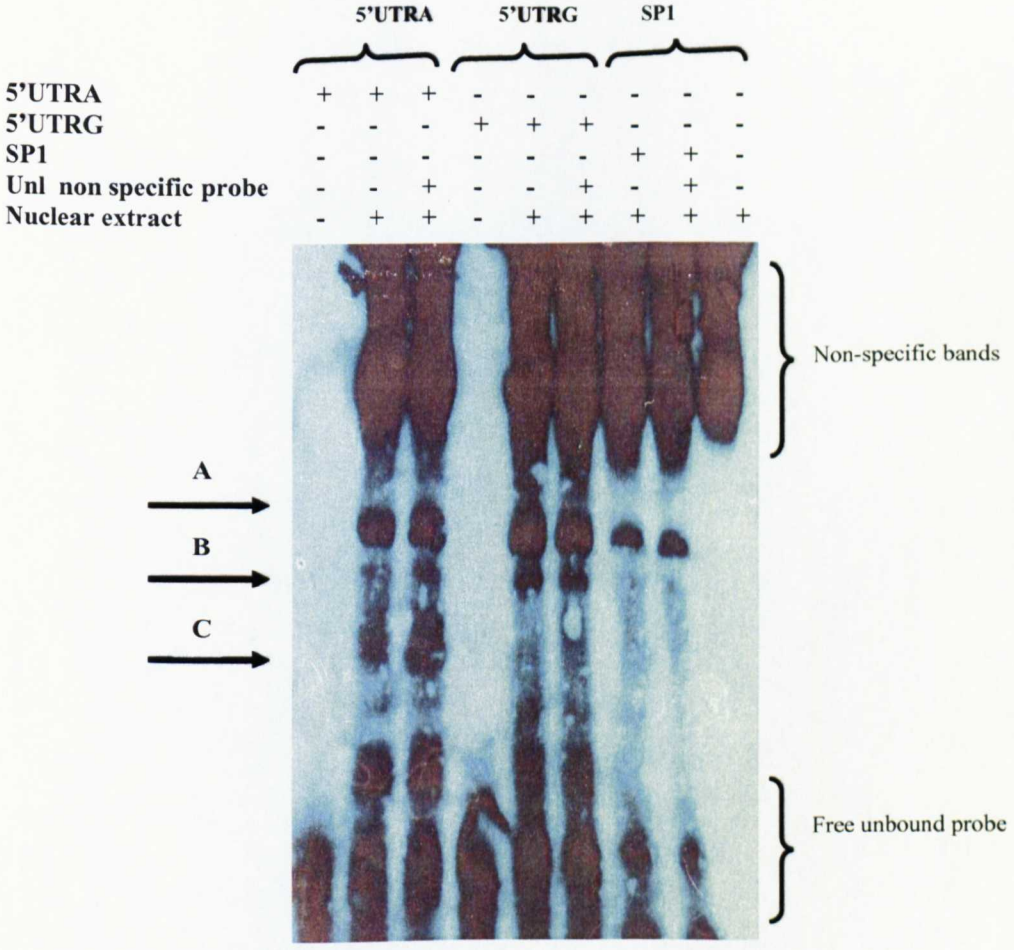
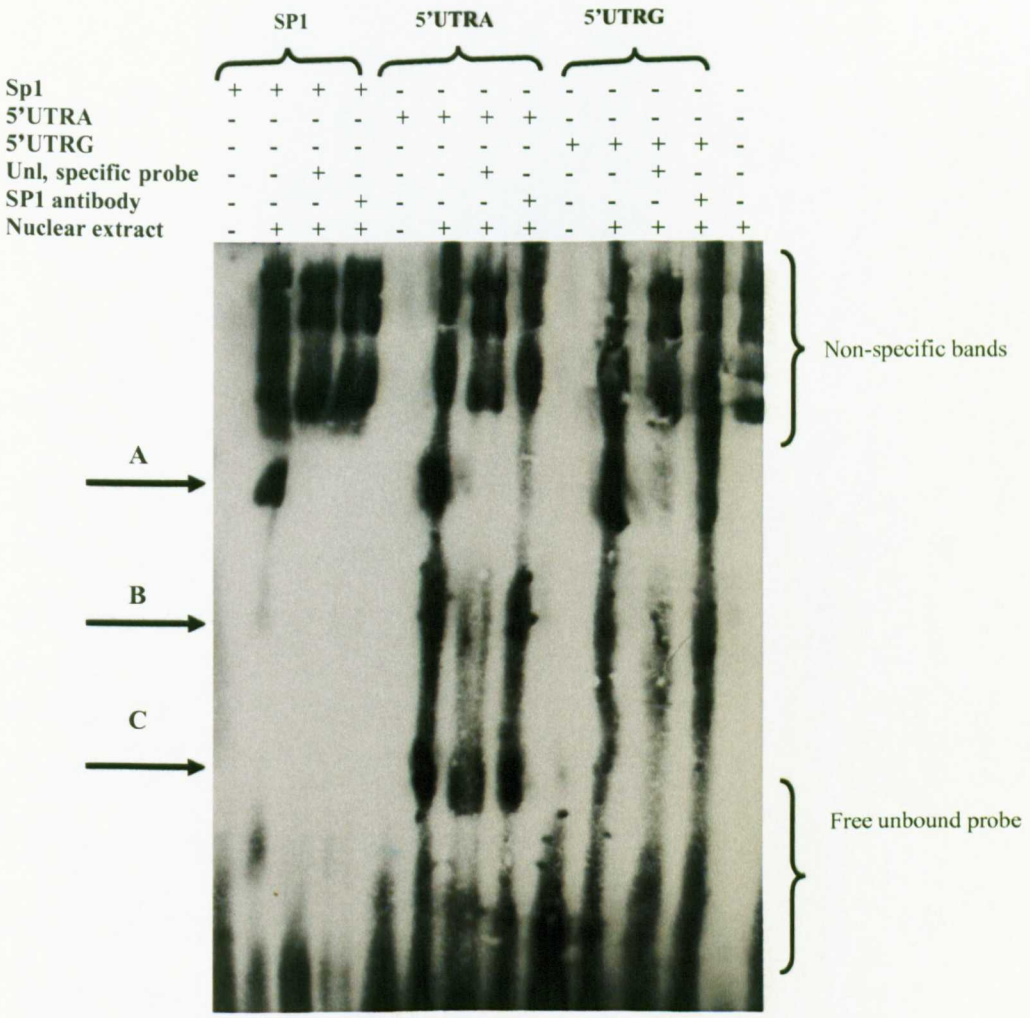


Figure 3-26: Supershift assays using antibodies to Sp1 with the 5'UTR of the EGF gene

The binding pattern is as described in Figure 3-27, again the bands are specific as they disappeared after addition of 100 molar excess of specific unlabeled probe. The arrow A indicates protein–DNA interactions observed using the three probes in the absence of the Sp1 antibody. This band significantly decreased when antibodies are included with 5'UTRA and Sp1 consensus site. In the case of 5'UTRG, the band is slightly diminished in the presence of Sp1 antibody.



3.5.5 Discussion

The results presented here showed the G allele of the SNP rs4444903 (c.61A>G) in the EGF-5'UTR region has greater transcriptional activity than the A allele in the absence of the EGFP in HepG2 (Table 3-9) and greater activity than the EGFP in JEG-3 cells (Table 3-10). When comparing the 5'UTR A and G independent from the EGFP we found that the G/A fold change was 2.6 at a P value of 0.004 in HepG2 cells over at least 5 experiments using different DNA preparations of the same construct. In the presence of the EGF promoter the G allele showed a 2.4 fold increase in the activity at P value of 0.02 in comparison to the EGF promoter in Jeg-3 cells. Hoogendoorn et al recommended three criteria for defining promoter haplotypes using reporter gene assays: 1) statistically significant differences in expression at $P<0.05$; 2) replication with independent construct preparations; 3) $\geq$ 1.5-fold difference in expression between haplotypes (Hoogendoorn et al., 2003). They indicated the last criterion to be the most important because it is the most consistently reproducible, did not fall in the range of random errors intrinsic to the assay and change of this magnitude when considered *in vivo* is equivalent to carrying an extra copy of the gene, comparing homozygotes for the high activity allele to homozygotes for the low activity allele (Hoogendoorn et al., 2003). The data presented in the current study indicated that this sequence variation in the 5'UTR region of the EGF gene possibly confers a promoter activity. These findings are in agreement with previous studies in other cell lines including peripheral white blood cells, Chinese hamster ovary cells and glioma cell line (Shahbazi et al., 2002, Vauleon et al., 2007, Costa et al., 2007) that showed greater transcriptional activity with

the G allele. Previous studies have tested this variant only in the context of the promoter region; this study has delineated the region responsible for the variation in transcriptional activity by testing the 5'UTR region on its own.

Although this SNP at position +61 is 30 bp away from what is known as a proximal promoter (up to bp +35), the G allele introduces promoter properties to this region. Transcription factor binding prediction demonstrated that the A>G transition introduces binding sites for 3 transcription factors (Sp1, P300, and AML-1a) all of which are known factors in regulation of expression of the growth factors genes and genes related to cell cycle. TF search also identified potential binding sites for another 3 TF which can bind to both the A and G alleles (Elk-1, Th1, and STATx).

Multiple reports ascribe putative functionality to the 5'UTR region of genes in transcriptional regulation more commonly than previously thought. 5'UTR have been implicated in mechanisms to repress transcription of the downstream genes through binding to transcription factors. For example, p53 suppresses the expression of bcl-2 through a p53 response element located in the 5'UTR of the bcl-2 gene (Miyashita et al., 1994). Similarly, the heat shock transcription factor 1 (HSF-1) like sequence in the murine TNF-α 5'UTR binds HSF-1 and is required for HSF-1-mediated transcriptional repression in the minimal mouse TNF-A promoter (Singh et al., 2002). Also Berardi et al (2003) identified a region in the cyclin D1 5'UTR between +75 to +138 that selectively represses cyclin D1 gene expression in aging cells (Berardi et al., 2003). These

investigations support the findings of this study that the 5'UTR region of the EGF gene contributes to gene expression through repression of transcription from the A allele.

Sp1 is the transcription factor potentially involved in this regulation. Numerous Sp1 binding sites are present in the human genome. As a constitutive transcription activator, Sp1 is associated with genes involved in almost all cellular functions. However, a growing body of evidence indicates Sp1 as a repressor of gene expression (Law et al., 2011, Bilsland et al., 2006, Mottet et al., 2009). Consequently, Sp1 can be a transcription activator or repressor depending on sequence specificity, and the spatial availability and the recruitment of different protein complexes (Li and Davie, 2010). This may also explain why the TF sites prediction was contrary to our findings in indentifying the G allele as holding the potential binding sites for Sp1 rather than the A allele. We have shown that Sp1 factor specifically bound the A allele of the 5'UTR region. The G allele binding pattern was specific, but the addition of the Sp1 antibody only slightly diminished the band, indicating that binding of Sp1 to the G allele may occur in a conformation that differs from the A allele. The probability that Sp1 can bind different sequences in the same promoter and exert activation and repression function is reported in mouse type 4 adenylyl cyclase. The promoter of this gene carries more than one Sp1 site, the 5' one is an activator and the middle is a repressor (Rui et al., 2008). The other explanation for the DNA: protein binding pattern observed with the G allele is the binding to a protein with similar size and structure to Sp1, possibly the long isoform of the Sp3 although this was not tested. It has been demonstrated that Sp3 is a stronger

transactivator of the p21 promoter than Sp1 (Sowa et al., 1999). These findings merit further investigation.

Sequence variation at rs4444903 did not show any effect on the efficacy of translation in HepG2 or Jeg-3 cell lines, or in a cell free environment. We found that the A>G transition does not add or remove any stem loops to the long highly structured mRNA 5'UTR region. However, both alleles require around 128 kcal/mol energy for the unwinding by the helicase eIF4A for translation of the mRNA to continue. This amount of energy is large enough to inhibit the translation process by the scanning ribosome. In addition, 3 uORF were recognized in the 5'UTR region of the EGF gene, although none of them are affected by this mutation. The presence of the uORF is known to interfere with the efficiency of translation, as the production of short peptide (uORF) may complex with the translation machinery and hinders the translation process. Also the uORF may lead to targeting its own mRNA to a nonsense-mediated mRNA decay. Interestingly it has been shown that the transcript carrying the G allele of SNP rs4444903 has more mRNA stability compared to the transcript carrying the A allele in the HepG2 cell line (Tanabe et al., 2008). These mechanisms for gene regulation are frequently encountered with genes involved in cellular development, cell cycle, and oncogenesis like TGF-ß1 and BRCA1 in breast cancer. This study has shown that the 5'UTR of the EGF reduces the efficiency of translation of downstream luciferase by about 70%.

3.6 The effect of SNP rs4444903 on EGF and EGFR expression in the placenta

3.6.1 Introduction

Epidermal growth factor receptor (EGFR) is a trans-membrane growth factor receptor belonging to the c-erbB family of receptor tyrosine kinases. EGFR is a ubiquitously expressed 170-kDa glycoprotein that consists of an extracellular receptor domain, a trans-membrane region, and an intracellular domain. The extracellular domain is formed of 622 AA with 12 N-terminal glycosylation sites, and a cysteine rich domain forming a groove for EGF binding. The trans-membrane domain, the hydrophobic part of the receptor, consists of 23 AA. The intracellular domain is formed of 542 AA that carries the tyrosine kinase activity, a domain for internalization and multiple sites for phosphorylation at tyrosine, serine and threonine residues (Boonstra et al., 1985).

Multiple ligands bind to and activate the EGFR, including EGF, transforming growth factor-α (TGF-α), amphiregulin, and heparin-binding EGF. Ligand binding of EGFR stimulates the proliferation of a wide variety of cell types. EGFR is a transducer protein implicated in many types of cancer, and has been used as a prognostic marker for many types of tumours.

3.6.2 Signal transduction by EGFR

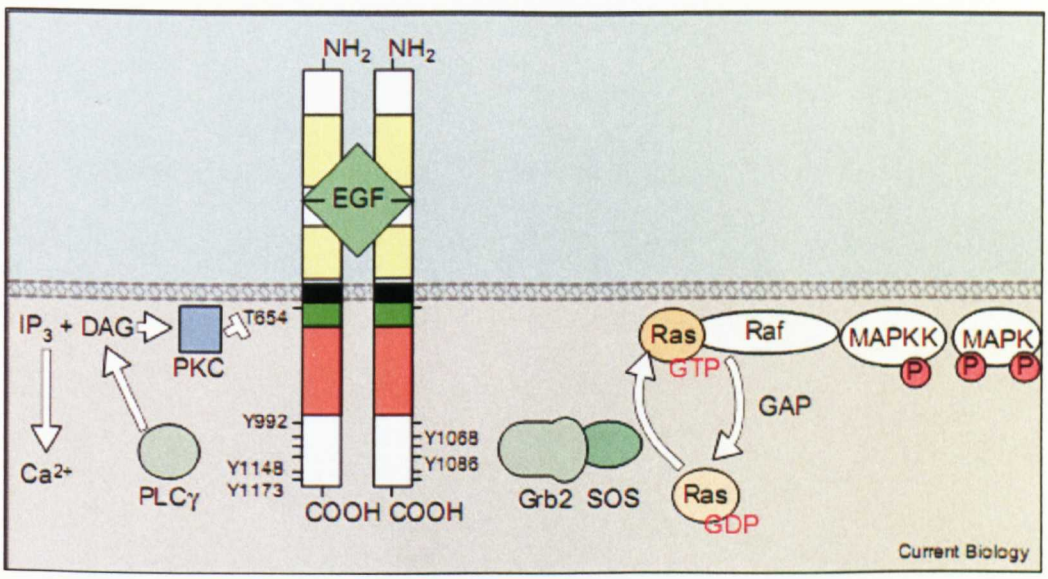
Ligand binding with EGFR results in dimerization of the receptor at the cell surface, followed by autophosphorylation of the intracytoplasmic domain, activation of the receptor tyrosine kinase domains as well as internalization of

the dimerized receptor. Phosphorylated tyrosine residues, located at the protein's carboxy terminus, serve as binding sites for the recruitment of signal transducers and activators of intracellular substrates, which then stimulate an intracellular signal transduction cascade (Sibley et al., 1987), (Figure 3-27).

Mitogen-activated protein kinase (MAP kinase) pathway is one of the downstream signalling pathways associated with EGFR. Upon ligand binding SOS (son of sevenless) guanine nucleotide exchange factor is recruited to the plasma membrane via the Grb2 adaptor protein. SOS stimulates the phosphorylation of the small G-protein Ras that subsequently stimulates the MAP kinase pathway to promote cell proliferation. A second pathway for EGFR signalling, the phospholipase C-γ (PLC-γ) transduction pathway, is involved mainly in stimulation of cell motility. phospholipase C-γ cleaves phospho-inositol biphosphate (PIP), with the production of diacylglycerol (DAG) and inositol 1,4,5-triphosphate (PI3) followed by activation of protein kinase C (PKC) and an increase in intracellular calcium. Activated protein kinase C phosphorylates T654 in the juxtamembrane part of the EGFR, which leads to disruption of the Ras–MAP kinase pathway and further biases the response to EGFR activation towards cell motility. In the mean time PI3-kinase blocks the activity of the inhibitory tyrosine protein phosphatase that is capable of terminating EGFR activation, implying a regulatory feedback loop between different signalling pathways for EGFR (Bogdan and Klambt, 2001).

Figure 3-27: The EGFR-signalling pathway

A simplified schematic showing the downstream signalling pathways involved in signal transduction. EGF: Epidermal Growth Factor, MAPK: mitogen activated protein kinases, MAPKK: MAP kinase kinase, PLC-γ: phospholipase C- γ, DAG: Diacylglycerol, IP3: Inositol1,4,5-triphosphate, PKC: Protein Kinase C, SOS: son of sevenless, Grb2: Growth factor receptor-bound protein 2 (Bogdan and Klambt, 2001), with permission.



3.6.3 Internalization and down regulation of the EGFR

Activation of the EGFR by its ligand leads to rapid internalization of the receptor through clathrin-coated pits. After internalization, EGF–receptor complexes appear consecutively in early endosomes and multivesicular bodies. During passage through the endosomal compartments, these complexes are either recycled back to the cell surface or sorted to late endosomes and lysosomes, where both ligand and EGFR are proteolytically degraded. The accelerated internalization and efficient lysosomal targeting of ligand-occupied receptors results in a down-regulation of the EGFR and the related responses. This is because rapid sequestration of the receptor into endosomes removes it from its

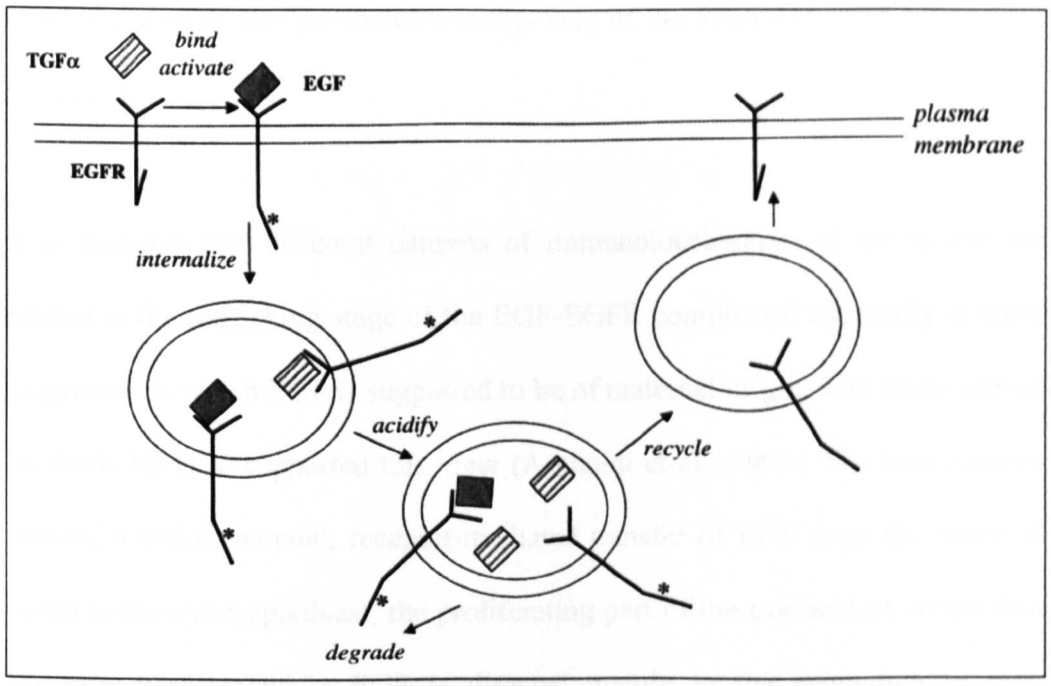
downstream target, Ras, the majority of which is constitutively associated with the plasma membrane (Wells et al., 1990, Wells, 1999), (Figure 3-28).

TGF-α, a highly potent EGFR agonist has an identical affinity constant for binding to the EGFR and similar internalization kinetics, but dissociates at a higher pH (half-maximal dissociation at pH 6.9). Targeting of the receptor for degradation in lysosomes (down regulation) is achieved more effectively by EGF than by TGF-α in murine B82L cells expressing the human EGFR or in human keratinocytes with endogenous EGFR. Also the refractory period after which the receptor can bind new ligands was 2.5 hours for TGF-α and 5 hours for EGF. The recovery was more sensitive to protein synthesis inhibitor in the case of EGF indicating that new EGFR synthesis was required. The greater biological potency of TGF-α may be due to the repeated presentation of recycled receptors for TGF-α and EGF at the cell surface, resulting in several rounds of signalling (Ebner and Derynck, 1991), (Figure 3-28).

Exogenous EGF was shown to have growth inhibitory effects on selected cell types that constitutively secrete TGF-α. For example, treatment of A-431 cells, which over expresses EGFR, with growth-inhibitory concentrations (10^{-9} M) of EGF leads to induction of the cell cycle inhibitor cyclin-interacting protein 1 (CIP1) with subsequent growth arrest. EGF at concentrations of (10^{-12} M) was found to be growth-stimulatory (Chajry et al., 1994).

Figure 3-28: Ligand-regulated internalization and degradation of EGFR

Upon ligand binding EGFR is activated and the receptor is internalized via clathrin-coated pits. This initial action is consistent between EGF and TGF-α. In the acidified endosomal compartment, EGF remains bound and directs the receptor and its bound EGF to degradation; TGF-α dissociates and the receptor is recycled while the free TGF-α is degraded (Wells, 1999), with permission.



3.6.4 EGF-EGFR in trophoblast (in normal and abnormal pregnancy)

The role of EGF in the developmental and functional aspects of the trophoblast is well defined. However reports of the spatial and temporal pattern of EGF and EGFR distribution in different trophoblast lineages are variable among researchers.

Maruo et al (1987) showed that EGFR is expressed more intensely in the 1st trimester ST trophoblast compared to mid and late term ST trophoblast, which

explained why EGF increases the secretion of hCG and hPL after a short period of stimulation in early compared to late trophoblast (Maruo et al., 1987). In contrast to this pattern of expression, EGFR is expressed more intensely in the proliferating CT trophoblast of early pregnancy, compared to ST trophoblast, regulating mainly the proliferative component of the trophoblast (Muhlhauser et al., 1993).

It is probable that different patterns of immunolocalization of the EGFR are related to the trafficking stage of the EGF-EGFR complexes, especially in early pregnancy in which EGF is suggested to be of maternal origin. The study carried by Arnholdt et al supported this view (Arnholdt et al., 1991). The investigators showed a trans-syncytial, receptor-mediated transfer of EGF from the maternal blood to the cytotrophoblast, the proliferating part of the trophoblast, in the first trimester placenta while EGF is directed mainly to the syncytium in term placenta. Also, Duello et al (1994) demonstrated that EGFR is localized to both trophoblast layers in first- and third-trimester placental in a pattern consistent with the activation and internalization of the EGFR. The investigators were also able to show an element of variation in the staining of adjacent syncytium, a variable which may explain conflicting reports regarding the pattern of EGFR in the trophoblast layers (Duello et al., 1994). Moreover, timing of sampling in each study can explain the variation in such a dynamic, rapidly growing and differentiated tissue.

Many studies have suggested the presence of an EGF/EGFR autocrine and paracrine mechanism in human trophoblast cells. Mauro et al (1992, 1995)

demonstrated the functional effects of EGF on trophoblast at 2 different stages, 4-5 and 6-12 weeks gestation. EGF stimulated proliferation in the early stage trophoblast as indicated by enhanced staining for Ki-67 antibody. In the late stage trophoblast, EGF induced differentiation of the trophoblastic explants as indicated by increased secretion of hCG and hPL on addition of EGF to culture media. Immunostaining localized EGF to the CT trophoblast and EGFR to the syncytiotrophoblast. The authors also demonstrated the secretion of EGF by the trophoblast tissue itself referring to the autocrine nature of this ligand receptor system (Maruo et al., 1992, Maruo et al., 1995). Amemiya et al (1994) showed the same findings in 6-9 weeks gestation trophoblast and confirmed the presence of EGF and EGFR by immunostaining and RT-PCR (Amemiya et al., 1994).

Many reports have shown disturbances of EGF/EGFR in abnormal pregnancy. Fujita et al (1991) showed reduced binding sites for EGF in the plasma membrane of term placental tissues in FGR compared to normal pregnancy. They were also able to show reduced mRNA levels of the EGFR, indicating down regulation of the receptors (Fujita et al., 1991). Another study showed reduced EGFR autophosphorylation and reduced tyrosine kinase activity in the plasma membrane fraction of the chorionic villi derived from placenta of FGR cases. The alteration of EGF-dependent tyrosine kinase activity appeared to be specific, as autophosphorylation of insulin receptors was preserved. Reduced EGF binding in the FGR placenta might explain the reduced receptor phosphorylation in this group of cases all of which presented with a maternal history of smoking and preeclampsia (Fondacci et al., 1994). Higher EGFR

concentrations in the placenta and fetal membrane in hypertensive pregnancy have also been reported (Ferrandina et al., 1995).

EGF and EGFR are expressed in the trophoblast throughout gestation to regulate its proliferation and differentiation. Intrauterine growth restriction and preeclampsia are both associated with various alterations of the placental EGF/EGFR system.

3.6.5 Aim of the study

This is an exploratory study to investigate the effects of genotype at rs4444903 in the 5'UTR of EGF on the expression of EGF in term placenta from healthy and preeclampsia-FGR complicated pregnancies.

3.6.6 Methods

3.6.6.1 Sample collection

Placental tissues were available from 9 healthy pregnancies and 16 complicated pregnancies (8 preeclampsia and 8 preeclampsia-FGR placentas). Preeclampsia was diagnosed according to ISSHP guidelines: systolic blood pressure of 140 mmHg or more and diastolic pressure of 90 mmHg or more on 2 occasions after 20 weeks gestation in a previously normotensive woman, and proteinuria exceeding 300 mg/L, or 2+ on dipstick analysis of midstream urine if 24-h collection result was not available (Brown et al., 2001). FGR was defined as corrected birth weight centile <10^{th} centile. These tissues were kindly provided by Dr Hiten Mistry; details about sample collection have been previously published (Mistry et al., 2010). In summary, these samples were taken from midway between the placental border and cord insertion site within 10 min after delivery. After membrane removal, the tissues were washed in ice cold 1X PBS to remove maternal blood contamination. Parts of the tissues were snap frozen in liquid nitrogen and kept at -80°C, while parts were paraffin embedded for immunohistochemistry. Informed consent was taken at the time of sample collection and the study was approved by the local ethics committee.

3.6.6.2 Immunohistochemical staining for EGFR and EGF

Serial sections of 5 μm thick formalin- fixed paraffin - embedded tissue were cut and mounted onto glass slides. After incubation at 60°C for 10 min, sections were deparaffinized by immersing in xylene (twice), and rehydrated through descending graded alcohol (100, 90, and 70%) 3 min each. Antigen retrieval was

carried out by Proteinase K (DAKO – Proteinase K ready to use – S3020) at 37°C for 10 min for EGFR, and by heat induced retrieval where sections were boiled in citrate buffer (0.1M concentration at PH 6.0) using a microwave for 20 min for EGF. The slides were loaded on Sequenza plates and placed in Sequenza trays then rinsed with TBS pH7.6. Endogenous hydrogen peroxidase activity was blocked by applying Dako Peroxidase block (DAKO REAL Peroxidase-Blocking solution–S2023) for 5 min. For blocking of nonspecific binding, sections were treated for 5 min with 200µl of UV block (Ultra V Block–ThermoScientific–TA-125-UB). Sections were incubated for 60 min at room temperature with optimized diluted primary antibody, made up in Dako antibody diluent (DAKO Antibody Diluent – 50809) (Table 3-11). A negative control was included by omitting incubation with the primary antibody. A positive control for each antibody was used to ensure specificity (Table 3-11).

Table 3-11: Details of antibodies and positive control tissues used in this study

The optimal dilution for each antibody was selected as the lowest dilution that gives maximal specific reactivity and minimal background staining

Antibody	Clone	Dilution	Control tissue	Antibody Supplier
Mouse anti-EGFr antibody – 28-0005.	31G7	1/60	Placenta	Zymed
Rabbit polyclonal anti-EGF – 07-1432	-	1/300	Tonsil	Millipore

After washing unbound primary antibody, immunostaining was performed using DAKO REAL EnVision Detection System, Peroxidase/DAB Rabbit/Mouse – code K5007, in which bound antibody was developed by using 3,3'-Diaminobenzidine tetrahydrochloride for 6 min at room temperature. Sections

were then counterstained with Mayer's haematoxylin, dehydrated with ascending graded alcohol, cleared in xylene, and mounted with DPX.

3.6.6.3 Analysis of the EGF and EGFR immunostaining of the placenta

A visual examination was performed to ensure specificity and distribution of immunolabelling. All slides were analysed by the same observer (Dr. Paula Williams) who was blinded to which slides were from cases and which were from controls. Five randomly selected fields were captured on NIS-Elements F2.20 (Nikon UK Ltd, Surrey, UK) at 200X magnification. The total percentage of positive labelled areas per 200X magnification field was determined using the positive pixel count function of the ImageScope (Aperio Technologies Ltd, Bristol, UK). Results are expressed as 'positivity' which takes into account both the number of positive pixels and the intensity of staining (Williams et al., 2010).

3.6.6.4 Genotyping of the placental samples

25 mg of frozen tissue was washed in ice cold PBS to eliminate any maternal blood which may interfere with the genotyping. The tissue was then subjected to homogenization in 80µl of PBS using 2 ml Lysing Matrix A tubes with granite matrix and ¼ ceramic spheres (www.qbiogene.com) for 40 sec in Fastprep®24 homogenizer (www.mpbio.com). DNA was extracted from the cell homogenate using QIAamp DNA Mini Kit (www.qiagen.com) according to the manufacturer's protocol. DNA was quantified and stored at -20^{0}C for genotyping. For genotyping of SNP rs4444903, 100 ng DNA was PCR amplified using primer pair EGF-PRO-S and EGF-PUTR-AS (Table 3-7) and the PCR conditions described previously (Section 3.5.3.1). PCR clean up was carried out

followed by DNA sequencing using 5pmol of EGF-PRO-S primer. DNA sequence was viewed by Chromas Lite, and aligned with ClustalW against the reference sequence.

3.6.6.5 Statistical analysis

Statistical analysis was performed using SPSS for Windows version 16.0. Summary data are presented as means (SEM) or number (frequency) as appropriate. Between-group comparisons were made using oneway ANOVA or Student's t test. The null hypothesis was rejected where $P < 0.05$.

3.6.7 Results

Mothers from healthy and preeclamptic-IUGR pregnancies were similar in age, parity, BMI, and placental weight distribution. By definition systolic and diastolic blood pressure were significantly higher in the preeclampsia group compared to healthy controls. Fetal weight and gestational age at delivery were lower in the preeclampsia group compared to healthy controls (Table 3-12).

Table 3-12: Demographic and pregnancy characteristics of subject groups

Data are presented as mean (SEM) except for parity and Caesarean section data presented as number (percentage).

	Healthy pregnancy (n=9)	Preeclampsia & FGR (n=16)	P value
Maternal age (years)	26.44 (2.11)	31.50 (1.55)	0.07
Nulliparous	5.0 (55.6)	10 (62.5)	0.73
BMI(kg/m2)	28.64 (1.88)	26.42 (1.24)	0.32
Mean systolic blood pressure (mmHg)	117.17 (1.21)	157.81 (1.61)	<0.001
Mean diastolic blood pressure (mmHg)	76.61(0.79)	97.81 (1.07)	<0.001
Gestational age at delivery (weeks)	39.46 (0.27)	36.58 (0.95)	0.01
Caesarean section	1.0 (11)	6 (37)	0.20
Fetal Weight (kg)	3.52 (0.10)	2.68 (0.25)	0.02
Birth weight centile	48.83 (8.16)	33.55 (10.10)	0.21
Placental wt (g)	660 (50.55)	602 (37.07)	0.36

Visual checks showed that EGFR staining was of moderate to high intensity and none of the slides showed < 75% staining. Staining for EGF was shown to be mild to moderate and none of the slides showed < 10% staining. EGFR staining was mainly localized to the trophoblast tissues, and was more intense in the ST trophoblast apical membrane. Although the CT trophoblast was relatively thinned out as expected in these term placental samples, EGFR immunoreactivity was apparent in the cytotrophoblast. Both membranous (inner membrane of the villous trophoblast) and cytoplasmic staining was evident (Figure 3-29). Some of the slides showed no EGFR staining of the fetal side of the cytotrophoblast.

EGF immunoreactivity was mainly cytoplasmic in nature. However some patchy staining of the maternal surface of the syncytium and some nuclear staining were observed. Also some positivity was detected in the cellular component of the villous stroma, around the fetal blood vessels, fibroblasts, and Hofbauer cells (villous macrophages), (Figure 3-30). We did not observe any variation in the EGFR or EGF staining of placental tissues from healthy (9 subjects) compared to preeclamptic-FGR (16 cases) pregnancies with P= 0.86 and 0.93 for EGF and EGFR respectively (Figure 3-31). EGFR and EGF immunostaining was similar in the placenta when the analysis was carried for healthy pregnancy compared to preeclampsia (8 cases) and preeclampsia-FGR placentas (8 cases). EGFR positivity was (0.42 ± 0.04, 0.43 ± 0.05, and 0.45 ± 0.02) in healthy, preeclampsia and preeclampsia-FGR placentas respectively (P=0.89). EGF positivity was (0.14 ± 0.03, 0.11 ± 0.03, and 0.13 ± 0.03) in healthy, preeclampsia and preeclampsia-FGR placentas respectively (P= 0.63). There was

no correlation between EGF and EGFR immunostaining of the placenta (correlation coefficient r = 0.21, P = 0.30).

Because no variation was detected between healthy and diseased pregnancies in immunostaining, placental staining for EGF and EGFR was analyzed as a function of genotype for cases and controls pooled together. The AA genotype showed the weakest staining for EGF compared to AG and GG genotype (Figure 3-34). In contrast the AA genotype was associated with stronger staining for EGFR compared to AG and GG genotype. These variations did not reach statistical significance when analysed by ANOVA across all genotypes (P= 0.18 and 0.60) for EGF and EGFR respectively (Figure 3-32).

Figure 3-29: EGFR staining of the placenta

EGFR expression is observed at non significant decreasing intensity from AA (A1), AG (A2), to GG (A3) placentas (P=0.6). Both layers of the trophoblast are expressing EGFR (green arrow); the staining is mainly localized to the cell membrane and less intense in the cytoplasm. The syncytiotrophoblast is more intensely staining especially at the maternal border (white arrow). A4 represents the negative control with omission of the primary antibody, A5 is the placenta considered as a positive control. 200X field for term placenta stained with Mouse anti-EGFR antibody (Zymed Cat # – 28-0005), 1/60 dilution, and developed by DAKO REAL EnVision Detection System, code K5007.

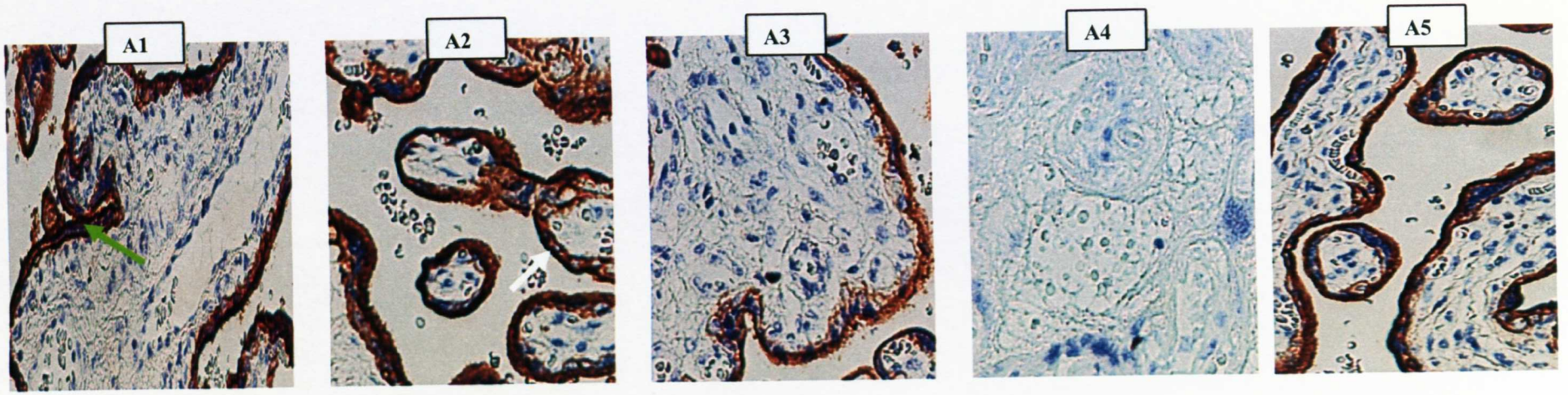


Figure 3-30: EGF staining of the placenta

EGF expression is observed at non significant increasing intensity from AA (B1), AG (B2), to GG (B3) placentas (P=0.18). Both layers of the trophoblast are expressing EGF, mainly cytoplasmic (Black arrow), but nuclear and membranous staining was also noticed (red arrow). EGF immunoreactivity is also noticed around fetal blood vessels (blue arrow) and in Hofbauer cells (yellow arrow). B4 represents the negative control with omission of the primary antibody, B5 is the tonsil considered as a positive control. 200X field for term placenta stained with rabbit polyclonal anti-EGF (Millipore Cat # 07-1432), 1/300 dilution, and developed by DAKO REAL EnVision Detection System, code K5007.

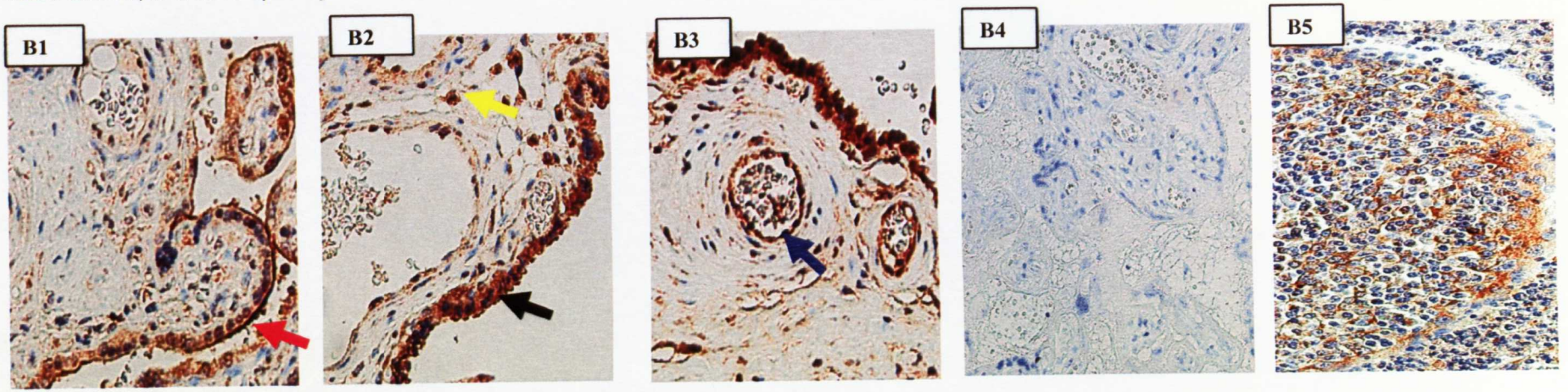


Figure 3-31: EGF and EGFR immunostaining in placentas from normal and preeclampsia-FGR pregnancies

Data shown as mean ± SEM of positivity/200x field.

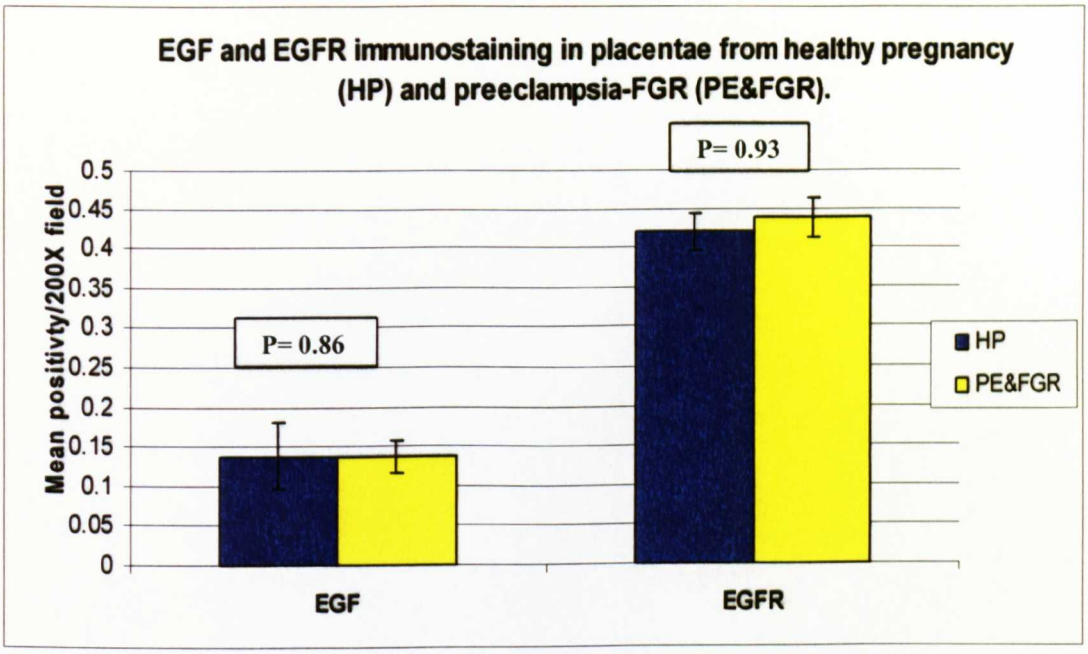
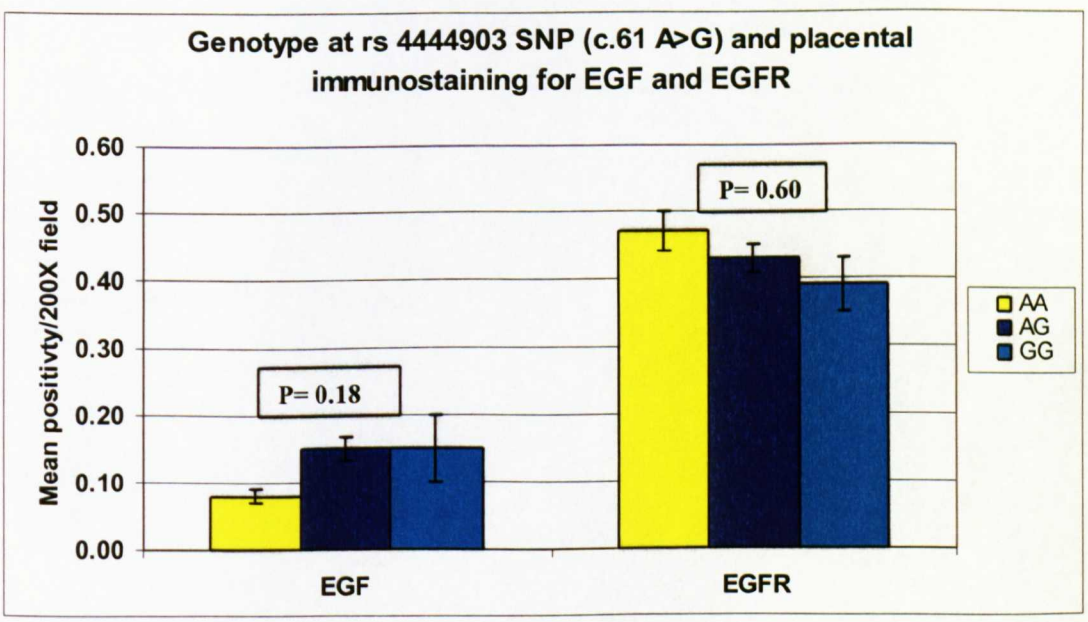


Figure 3-32: EGF and EGFR immunostaining in placentae according to genotype at marker rs4444903 of the EGF gene

Data are shown as mean ± SEM of positivity/200x field. Genotype numbers were 5, 16, 4 for AA, AG and GG genotypes respectively.



3.6.8 Discussion

In this study showed that EGFR expression in the placenta was mainly localized to the ST trophoblast (maternal side) and less intensely to the inner surface of the CT trophoblast (fetal side) towards the villous stroma. Some areas showed positivity in the cytoplasm of both layers. This pattern was in agreement with previous studies (Duello et al., 1994). Rao et al showed that the maternal interface of the ST trophoblast holds 40% more available receptors for binding to radiolabelled EGF compared to the basolateral interface (fetal) (Rao et al., 1995). Some patchy staining of the inner membrane of the CT trophoblast was observed, which is consistent with synthesis of EGF by CT trophoblast and its transfer to the syncytium during the fusion process (Duello et al., 1994) or with transsyncytial transfer of the EGF bound EGFR from the maternal side to the fetal side (Arnholdt et al., 1991). This fits well with the hypothesis of receptor internalization upon EGF binding.

The pattern and the low intensity of EGF staining of the term placenta are in agreement with previous studies (Hofmann et al., 1991, Ladines-Llave et al., 1991). The lower levels of expression of EGF in the placenta were also reported by Bissonnette et al (1992) who reported lower EGF mRNA and protein in term placenta (Bissonnette et al., 1992). Bass et al (1994) also reported very low to no expression of EGF mRNA or EGF protein in the culture media of cytotrophoblast (Bass et al., 1994). These findings support the hypothesis that EGF at the fetomaternal interface is mainly of maternal origin.

This study included placentas from 9 healthy pregnancies and 16 preeclampsia-FGR cases, and EGF and EGFR expression did not differ between the two

groups, as determined by immunostaining. The data regarding the EGFR levels in the pathological placenta are conflicting. In agreement with these data, Calvo et al (2004) demonstrated no variation in the EGFR mRNA levels between appropriate and small for gestational age placentas (Calvo et al., 2004. Reduced EGFR activity and mRNA expression in the placenta from FGR cases were reported {Fondacci, 1994 #156, Fujita et al., 1991), both studies were comparable to this study in terms of sample size. Although EGF was immunolocalized in the placenta no data are available to describe variation in the pathological placenta. The small sample size in the present investigation and previous studies, incomplete data on the mode of delivery, and variation in methods of detection hamper efforts to define possible alterations in the EGF/EGFR system in the placentas of abnormal pregnancies, and none of the findings can be considered conclusive. The present study used placenta collected during labour, which has been reported to result in up regulation of the EGFR (Gargiulo et al., 1997). This may mask any variation between normal and pathological placentas. Furthermore, immunohistochemical studies have shown that staining intensity of EGF and EGFR diminishes with increasing gestational age (Arnholdt et al., 1991, Duello et al., 1994). Moreover, it is well recognized that the pathological changes related to preeclampsia originate in early placenta, so it would be interesting to investigate early placenta tissue in that regard. It is difficult to obtain early placental tissues from pregnancies which known to be destined to develop preeclampsia after 20 weeks of gestation. Chorionic villus samples obtained for the purpose of genetic counselling in pregnancies which are subsequently affected by pre-eclampsia may provide limited amounts of suitable tissues, but potential complications and ethical issues should be considered.

Previous studies that measured EGF levels in relation to genotype at SNP rs4444903 in glioblastoma (Bhowmick et al., 2004), colorectal cancer (Spindler et al., 2007) and hepatocellular carcinoma (Tanabe et al., 2008) showed that the G allele at SNP rs4444903 was linked to higher EGF expression both at the mRNA and protein levels. However, in the current study, the G allele in the placenta was not associated with higher EGF expression compared to the A allele, (Figure 3-32). Likewise the levels of EGFR expression tended to be lower in placentae with the GG or AG genotype compared to the AA genotype but the variation was not significant (P 0.34). Bhowmick et al (2004) also noted that the GG genotype was associated with non-significant lower expression of EGFR in glioblastoma tissue. In trophoblast cell culture, EGF at low concentrations (0.1 to 1 nM) stimulated receptor biosynthesis, an effect that diminished at higher EGF concentrations. The effect of EGF on receptor degradation was apparent only at saturating concentrations (> 10 nM), (DePalo and Das, 1988). The situation in vivo seems different, as the GG genotype at this locus does not lead to higher EGF production to the critical values that downregulate the EGFR. However the sample size that has been used in this study is small and conclusive data about EGF expression in the placenta as a function of genotype at rs4444903 require further verification in a larger sample set, preferably in earlier gestational age placentas

3.7 Conclusion

When single nucleotide polymorphisms are found to be associated with a complex disorder it is vital to identify whether and how this SNP affects gene function, providing further evidence that this gene is implicated in the disease and giving a better understanding of the disease process. This study has shown no role for SNP rs4444903 of the 5'UTR region of the EGF gene in the risk of preeclampsia or FGR. Functionality of this polymorphism has been investigated in hepatocellular carcinoma and choriocarcinoma cell lines. The higher transcriptional activity of the G allele is possibly related to a repressor effect exerted by Sp1 transcription factor on the A allele. However an interaction of the G allele with Sp1 transcription factor was specific and merits more investigation. This study showed no functional impact for this SNP on EGF or EGFR expression in a small study of term placental tissues. This study excluded an effect of SNP rs4444903 on the translation efficiency of the downstream gene in HepG2, Jeg-3 cells and in a cell-free environment. SNP rs2237051 does not disrupt a functioning exonic splicing enhancer in EGF exon 14 in SGH-PL4 and HepG2 cells as shown by the minigene assay.

4 Final discussion and future work

4.1 Summary

There are two main areas of work presented in this thesis. Firstly, as there is a known relationship between diabetes and preeclampsia/FGR, maternal genes which are known to be reproducibly associated with T2D (TCF7L2, FTO, PPAR-γ, CDKN2B-AS1 and KCNJ11) were tested for association with preeclampsia and fetal growth restriction. We also sought to investigate whether these genes contribute to the variation in birth weight. Secondly, the role of EGF (a major regulator of the placentation process) in susceptibility to preeclampsia and FGR, and regulation of birth weight was also investigated. In addition, molecular genetic studies were carried out to investigate the functional effects of the SNPs rs4444903 and rs2237051 on the expression of EGF *in vitro* and *in vivo*.

This study showed for the first time that a maternal haplotype on the T2D and CAD region of the CDKN2B-AS1 gene on chromosome 9p21 was a risk variant for growth restricted pregnancies. The level of significance for the association of CDKN2B-AS1 haplotype and the risk of fetal growth restriction on global haplotype testing was P=0.005; comparing the C-G (risk haplotype) to the T-G (common or the reference haplotype) the level of statistical significance was P=0.0008. This was robust enough to withstand correction for multiple testing. No previous studies in humans relate CDKN2B-AS1 to FGR, but animal studies have noted that deletion of the orthologous region on mouse chromosome 4 was associated with higher offspring birth weight (Visel et al., 2010), suggesting a role for CDKN2B-AS1 in growth regulation.

The other 9 investigated SNPs in TCF7L2, FTO, PPAR-γ, and KCNJ11 showed no association with the risk of having growth restricted pregnancies, nor were any of the loci investigated associated with birth weight.

None of the T2D genes was associated with preeclampsia risk. This study confirmed the findings of two previous studies (Laasanen et al., 2002, Wiedemann et al., 2010) that SNP rs1801282 of PPAR-γ is not associated with preeclampsia. In contrast to the findings in a Finnish population (Peterson et al submitted for publication) that variants in the CDKN2B-AS1 gene were associated with an increased risk of preeclampsia, no association was found in this study, possibly due to differences in the LD pattern of the chromosome 9p21 region in UK and Finnish populations. Other possibilities are that this study is underpowered to detect the association, or the findings of the Finnish study were spurious. None of the other variants of T2D genes was associated with preeclampsia risk. Consequently, the case control study did not confirm a shared genetic basis between T2D and preeclampsia.

The G allele of the SNP rs4444903 of the 5'UTR of the EGF gene showed no effect on preeclampsia or FGR risk, likewise the A allele at the SNP rs2237051 showed no effect on either the risk of preeclampsia or FGR. So the study reported no genetic evidence for the major role of maternally derived EGF on the placentation process early in pregnancy (Maruo et al., 1992, Morrish et al., 1997). The rare genotype GG at SNP rs4444903 was associated with the highest systolic blood pressure measures in healthy pregnant mothers. The effect on blood pressure is thought to be related to the systemic effect of maternal EGF on

vascular tone (Lundstam et al., 2007). In relation to birth weight, the rare genotype at these 2 SNP of the EGF gene was associated with the lowest birth weight in healthy babies though the variation in birth weight was not statistically significant. In view of a previous study which showed that the haplotype comprised of the rare alleles (G-A) at these loci was transmitted more frequently to FGR babies from heterozygous parents (Dissanayake et al., 2007), it is possible that it is the fetal not maternal genotypes at these loci that contributes to the variation in birth weight.

The G allele of the SNP rs4444903 in the 5'UTR of the EGF gene showed higher transcriptional activity in hepatocellularcarcinoma (HepG2) cell lines, but no evidence that allelic variation affects translational efficiency. The higher transcriptional activity of the G allele was also evident in glioma (Bhowmick et al., 2004), and colon cancer (Spindler et al., 2007). The study showed that this transcriptional activity resides in the 5'UTR and is independent of the EGF promoter. DNA-protein interaction with the transcription factor Sp1 that appears to impose transcriptional repression on the A allele is offered as an explanation for the higher activity of the G allele, although these results are preliminary and require further exploration. No significant effect has been demonstrated for the variant rs4444903 on the EGF or EGFR expression of the term placental tissues.

4.2 Strengths and limitations of the study

1. Case-control study

The case-control study included women who were self identified as white western European women, minimizing population stratification that may confound the results. In addition to the diagnostic criteria used to define preeclampsia and FGR, there were some variations in phenotypic data between cases and controls, which is unavoidable in retrospective studies. For the pre-eclampsia study, differences between cases and controls were observed for BMI, the incidence of multiple pregnancy and primiparity. For the FGR study, differences between cases and controls were evident for BMI, smoking habits and chronic renal disease. All of these factors were adjusted for in the course of the statistical analysis, which imposed some reduction in the study power. Importantly, and in contrast to previous studies, this study included women with medical risk factors for preeclampsia, to avoid selecting against genetic factors shared with preeclampsia.

2. Molecular Studies

Although reporter gene assays are widely used to identify the effect of a polymorphism on gene expression, they have limitations as they may not reflect the *in vivo* situation (Cirulli and Goldstein, 2007). These limitations are related to gene context, the experimental technique used, and the use of cell lines.

Chromatin structure, and the presence of other regulatory elements that may be located up to 1 Mb from the gene promoters are not considered in the reporter gene system, so reporter gene assays may not mirror the *in vivo* transcription for a given gene (Buckland et al., 2005). In relation to the experimental technique it

has been shown that factors like the amount of DNA, the ratio between the test and control DNA, the quality and conformation of DNA used for transfection, the method of transfection, and the passage number of the host cells used in transfections can affect SNP-specific regulation in reporter gene assays (Karimi et al., 2009). Bearing this in mind, we used the minimal amount of DNA that can give consistent results; fresh preparations of DNA were used over experiments. The effect of the passage number of the cell on the absolute luciferase output but not the relative luciferase activity was indeed observed with Jeg-3, consequently after passage number 5, fresh frozen cells were brought to culture and used for transfection. However, the use of transient transfection methods in which multiple copies of the test DNA are expressed from extrachromosomal plasmid, may exhaust the regulatory protein machinery of the cells. Karimi et al (2009) suggested the use of stable transfection with the use of a single site integration method in which a single copy of the test DNA and the reporter gene are integrated at exactly the same genomic location and orientation in all cells within a chromatin context that is very similar to the genome *in vivo*. Obviously this method will be difficult to apply to the large number of constructs that have been tested in the current study, because creating such constructs and producing stable cell lines would require a considerable amount of time. It is for this reason that transient transfection methods are still widely used for experiments such as those described here.

Cell lines have been extensively used to investigate SNPs in regulatory regions on gene expression. The results from cell lines should be interpreted cautiously as the gene expression pattern may be altered in immortalised cell lines. Not only cells with monoclonal gene expression, unrepresentative of the polyclonal cell

populations, but also the growth rate and pattern, and the molecular repertoire of the cells do not reflect the *in vivo* situation (Gimelbrant et al., 2007).

There is no doubt that the use of immunohistochemistry technique is easy, cheap, reproducible, and can be applied to large number of samples at the same time. The most important issue in its superiority to Western blotting in protein detection is the localization of the protein of interest within the tissue under investigation, which of course gives an idea about its functionality. In the placental tissues used in this study, it was possible to locate the EGF and EGFR in different components of the trophoblast villi. But the IHC technique also has limitations as a quantitative technique: in this study this was addressed by the use of a computer based scoring system that can score a wide range of variability and is more objective than the manual scoring system. IHC also has limitations regarding tissue processing and time lapse between slide preparation and staining that may reduce the antigenicity of the protein and the staining outcome. It is worth noting that the slides used in the current study were freshly cut from the paraffin blocks just before the staining process. It is also important to recognise that EGF of maternal origin may be detected in placental tissues following internalization subsequent to binding to EGFR expressed on the syncytiotrophoblast.

4.3 Further work

Many studies have investigated the effect of EGF on the proliferative and invasive functions of the trophoblast, but in view of the findings of this study it would be interesting to test these roles under allelic effects on primary trophoblast cells or placental explants. Other approaches which would complement the findings of this study include an investigation of allelic imbalance in gene expression from endogenous heterozygous tissues. Quantitative PCR of EGF mRNA would complement the immunohistochemical results which suggest that EGF protein expression is associated with genotype.

There is compelling evidence that in genetic studies of complex disorders larger sample sizes generate more robust signals, and increase the probability of detecting variants with small effect sizes. A larger sample size can also enable analysis of subgroups, such as early onset and late onset preeclampsia which are thought to have different pathological mechanisms, and preeclampsia complicated by FGR. It would be valuable to know the glycaemic status of the mother to be able to delineate the genetic from environmental effects when testing the maternal T2D susceptibility genes and the risk for FGR. In studies like this which attempt to include mothers with known medical risk factors for preeclampsia an optimal design would include matching for these factors in cases and controls.

Such large and detailed genetic studies require collaboration between research groups. There are now a number of research groups within the international research community which encourage collaborative research. An example is the Norwegian mother and baby collection, "MoBa" (Magnus et al., 2006). DNA

samples from mother-father-baby triads are available from this collection of over 100,000 pregnancies for investigations of the causes of disease in mother and child. Such a collection would be valuable for the investigation of the role of fetal genotypes, and the interaction between fetal and maternal genotypes, as well as gene- environment interactions. Meta-analysis of the results generated by the current study and similar investigations which recruit participants regardless of their medical background risk, will enable the identification of mothers at high risk for preeclampsia, and related disorders, including eclampsia, FGR, and intrauterine death. In complex genetic disorders like T2D, CAD, preeclampsia, and FGR, genetic risk factors are interacting and modifiable by environmental factors. Understanding these factors will help to inform strategies directing life style modification, creating a public health orientation to lower the risk of the disease, or its complications.

The results from the first GWAS in preeclampsia carried out in UK as a part of WTCCC3 are expected shortly (https://www.wtccc.org.uk/ccc3/projects/ccc3_eclampsia.shtml). The study involves 2000 white UK women with preeclampsia who are otherwise healthy, with no known medical risk factors for preeclampsia. 6000 subjects from the UK 1958 birth cohort, with available GWAS data, will be used as the reference group. Results from studies such as this will offer new insights into the genetics and pathophysiology of preeclampsia. T2D provides a clear example of the potential of GWAS in complex disorders. Before the era of GWAS there were only 3 proven T2D susceptibility genes, PPAR-G, KCNJ11, and TCF7L2. Currently at least 35 loci are implicated in the risk of T2D, most of which would not previously have been considered on a biological basis as risk factors for T2D (Bonnefond et al., 2010). Recruiting adequate numbers of

women for GWAS of preeclampsia and FGR in those with pre-existing health problems would be challenging, though not impossible. A more practical approach would be to regard signals which arise in the course of GWAS of low risk pregnant women as priorities for studying as candidate genes in a high risk group. The emergence of new generation sequencing, with the capability of rapid, high throughput sequencing, makes whole genome sequencing available, exemplified by the 1000 Genomes Project. Whole genome sequencing has the advantage of revealing all genetic variation, SNP and non-SNP (eg. insertion deletion, and copy number variations), common and rare variations. Genome wide sequencing will augment the design of the variants used for GWAS by including rare variants. This of course will refine the search for the causal variants of disease. However the cost and the analytical difficulties may mean that using signals generated from the current GWAS platforms followed by deep sequencing is a more feasible strategy for future genetic studies.

The identification of susceptibility loci must be supplemented by functional studies to elucidate how these loci modify disease risk. The studies reported in this thesis include a wide range of strategies to determine functional effects of genetic variants mapped to regulatory regions, ranging from the use of bioinformatics to reporter gene assays, protein DNA interaction, and *in vivo* studies using human tissues. Advances in understanding the molecular genetics of complex disorders will help in understanding the pathological processes of the diseases. This will identify novel targets for the prediction, diagnosis, treatment, and importantly prevention of complex disorders, and raises the prospect of "personalized medicine" based on genetic variation information.

Currently genetic association studies of preeclampsia and FGR are mainly limited to populations of European descent. Extending these studies to other ethnic groups, for example populations of African origin who have shorter LD blocks, can help identification of causal variants. The author of this thesis has a personal interest in hoping that studies of the genetics of complex disorders of pregnancy will include the Egyptian population, at a level that engages the scientific community and leads to improvements in public health.

5 References

AAGAARD-TILLERY, K. M., PORTER, T. F., LANE, R. H., VARNER, M. W. & LACOURSIERE, D. Y. 2008. In utero tobacco exposure is associated with modified effects of maternal factors on fetal growth. *Am J Obstet Gynecol,* 198, 66 e1-6.

AL-ATTAR, S. A., POLLEX, R. L., BAN, M. R., YOUNG, T. K., BJERREGAARD, P., ANAND, S. S., YUSUF, S., ZINMAN, B., HARRIS, S. B., HANLEY, A. J., CONNELLY, P. W., HUFF, M. W. & HEGELE, R. A. 2008. Association between the FTO rs9939609 polymorphism and the metabolic syndrome in a non-Caucasian multi-ethnic sample. *Cardiovasc Diabetol,* 7, 5.

ALTSHULER, D. 2000. The common PPAR[ggr] Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nature Genet.,* 26, 76-80.

AMEMIYA, K., KURACHI, H., ADACHI, H., MORISHIGE, K. I., ADACHI, K., IMAI, T. & MIYAKE, A. 1994. Involvement of epidermal growth factor (EGF)/EGF receptor autocrine and paracrine mechanism in human trophoblast cells: functional differentiation in vitro. *J Endocrinol,* 143, 291-301.

ANANTH, C. V., KAMINSKY, L., GETAHUN, D., KIRBY, R. S. & VINTZILEOS, A. M. 2009. Recurrence of fetal growth restriction in singleton and twin gestations. *J Matern Fetal Neonatal Med,* 22, 654-61.

ANANTH, C. V. & VINTZILEOS, A. M. 2009. Distinguishing pathological from constitutional small for gestational age births in population-based studies. *Early Hum Dev,* 85, 653-8.

ANDERSSON, E. A., PILGAARD, K., PISINGER, C., HARDER, M. N., GRARUP, N., FAERCH, K., POULSEN, P., WITTE, D. R., JORGENSEN, T., VAAG, A., HANSEN, T. & PEDERSEN, O. 2010. Type 2 diabetes risk alleles near ADCY5, CDKAL1 and HHEX-IDE are associated with reduced birthweight. *Diabetologia,* 53, 1908-16.

ANDREASEN, C. H., STENDER-PETERSEN, K. L., MOGENSEN, M. S., TOREKOV, S. S., WEGNER, L., ANDERSEN, G., NIELSEN, A. L., ALBRECHTSEN, A., BORCH-JOHNSEN, K., RASMUSSEN, S. S., CLAUSEN, J. O., SANDBAEK, A., LAURITZEN, T., HANSEN, L., JORGENSEN, T., PEDERSEN, O. & HANSEN, T. 2008. Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes,* 57, 95-101.

ARNGRIMSSON, R. 2005. Epigenetics of hypertension in pregnancy. *Nat Genet,* 37, 460-461.

ARNGRIMSSON, R., BJORNSSON, S., GEIRSSON, R. T., BJORNSSON, H., WALKER, J. J. & SNAEDAL, G. 1990. Genetic and familial predisposition to eclampsia and pre-eclampsia in a defined population. *Br J Obstet Gynaecol,* 97, 762-9.

ARNGRIMSSON, R., HAYWARD, C., NADAUD, S., BALDURSDOTTIR, A., WALKER, J. J., LISTON, W. A., BJARNADOTTIR, R. I., BROCK, D. J., GEIRSSON, R. T., CONNOR, J. M. & SOUBRIER, F. 1997. Evidence for a familial pregnancy-induced hypertension locus in the eNOS-gene region. *Am J Hum Genet,* 61, 354-62.

ARNGRIMSSON, R., PURANDARE, S., CONNOR, M., WALKER, J. J., BJORNSSON, S., SOUBRIER, F., KOTELEVTSEV, Y. V., GEIRSSON, R. T. & BJORNSSON, H. 1993. Angiotensinogen: a candidate gene involved in preeclampsia? *Nat Genet,* 4, 114-5.

ARNGRIMSSON, R., SIGURARD TTIR, S., FRIGGE, M. L., BJARNAD TTIR, R. I., JONSSON, T., STEFANSSON, H., BALDURSDOTTIR, A., EINARSDOTTIR, A. S., PALSSON, B., SNORRADOTTIR, S., LACHMEIJER, A. M., NICOLAE, D., KONG, A., BRAGASON, B. T., GULCHER, J. R., GEIRSSON, R. T. & STEFANSSON, K. 1999. A genome-wide scan reveals a maternal susceptibility locus for pre-eclampsia on chromosome 2p13. *Hum Mol Genet,* 8, 1799-805.

ARNHOLDT, H., DIEBOLD, J., KUHLMANN, B. & LOHRS, U. 1991. Receptor-mediated processing of epidermal growth factor in the trophoblast of the human placenta. *Virchows Arch B Cell Pathol Incl Mol Pathol,* 61, 75-80.

ARROYO, J. A. & WINN, V. D. 2008. Vasculogenesis and angiogenesis in the IUGR placenta. *Semin Perinatol,* 32, 172-7.

ARYA, R., DEMERATH, E., JENKINSON, C. P., GORING, H. H. H., PUPPALA, S., FAROOK, V., FOWLER, S., SCHNEIDER, J., GRANATO, R., RESENDEZ, R. G., DYER, T. D., COLE, S. A., ALMASY, L., COMUZZIE, A. G., SIERVOGEL, R. M., BRADSHAW, B., DEFRONZO, R. A., MACCLUER, J., STERN, M. P., TOWNE, B., BLANGERO, J. & DUGGIRALA, R. 2006. A quantitative trait locus (QTL) on chromosome 6q influences birth weight in two independent family studies. *Hum. Mol. Genet.,* 15, 1569-1579.

ASHTON, S. V., WHITLEY, G. S., DASH, P. R., WAREING, M., CROCKER, I. P., BAKER, P. N. & CARTWRIGHT, J. E. 2005. Uterine spiral artery remodeling involves endothelial apoptosis induced by extravillous trophoblasts through Fas/FasL interactions. *Arterioscler Thromb Vasc Biol,* 25, 102-8.

ATAMER, Y., KOCYIGIT, Y., YOKUS, B., ATAMER, A. & ERDEN, A. C. 2005. Lipid peroxidation, antioxidant defense, status of trace metals and leptin levels in preeclampsia. *Eur J Obstet Gynecol Reprod Biol,* 119, 60-6.

AVERY, C. L., FREEDMAN, B. I., KRAJA, A. T., BORECKI, I. B., MILLER, M. B., PANKOW, J. S., ARNETT, D., LEWIS, C. E., MYERS, R. H., HUNT, S. C. & NORTH, K. E. 2006. Genotype-by-sex interaction in the aetiology of type 2 diabetes mellitus: support for sex-specific quantitative trait loci in Hypertension Genetic Epidemiology Network participants. *Diabetologia,* 49, 2329-36.

AYDIN, S., BENIAN, A., MADAZLI, R., ULUDAG, S., UZUN, H. & KAYA, S. 2004. Plasma malondialdehyde, superoxide dismutase, sE-selectin, fibronectin, endothelin-1 and nitric oxide levels in women with preeclampsia. *Eur J Obstet Gynecol Reprod Biol,* 113, 21-5.

BAETEN, J. M., BUKUSI, E. A. & LAMBE, M. 2001. Pregnancy complications and outcomes among overweight and obese nulliparous women. *Am J Public Health,* 91, 436-40.

BAKKETEIG, L. S., JACOBSEN, G., HOFFMAN, H. J., LINDMARK, G., BERGSJO, P., MOLNE, K. & RODSTEN, J. 1993. Pre-pregnancy risk factors of small-for-gestational age births among parous women in Scandinavia. *Acta Obstet Gynecol Scand,* 72, 273-9.

BALDING, D. J. 2006. A tutorial on statistical methods for population association studies. *Nat Rev Genet,* 7, 781-791.

BARKER, D. J. 1995. The fetal and infant origins of disease. *Eur J Clin Invest,* 25, 457-63.

BARKER, D. J. 2002. Fetal programming of coronary heart disease. *Trends Endocrinol Metab,* 13, 364-8.

BARKER, D. J., ERIKSSON, J. G., FORSEN, T. & OSMOND, C. 2002. Fetal origins of adult disease: strength of effects and biological basis. *Int J Epidemiol,* 31, 1235-9.

BARKER, D. J., HALES, C. N., FALL, C. H., OSMOND, C., PHIPPS, K. & CLARK, P. M. 1993. Type 2 (non-insulin-dependent) diabetes mellitus, hypertension and hyperlipidaemia (syndrome X): relation to reduced fetal growth. *Diabetologia,* 36, 62-7.

BARKER, D. J., OSMOND, C., GOLDING, J., KUH, D. & WADSWORTH, M. E. 1989a. Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. *BMJ,* 298, 564-7.

BARKER, D. J., WINTER, P. D., OSMOND, C., MARGETTS, B. & SIMMONDS, S. J. 1989b. Weight in infancy and death from ischaemic heart disease. *Lancet,* 2, 577-80.

BARROSO, I., GURNELL, M., CROWLEY, V. E., AGOSTINI, M., SCHWABE, J. W., SOOS, M. A., MASLEN, G. L., WILLIAMS, T. D., LEWIS, H., SCHAFER, A. J., CHATTERJEE, V. K. & O'RAHILLY, S. 1999. Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature,* 402, 880-3.

BASS, K. E., MORRISH, D., ROTH, I., BHARDWAJ, D., TAYLOR, R., ZHOU, Y. & FISHER, S. J. 1994. Human cytotrophoblast invasion is up-regulated by epidermal growth factor: evidence that paracrine factors modify this process. *Dev Biol,* 164, 550-61.

BASSOLS, J., PRATS-PUIG, A., VAZQUEZ-RUIZ, M., GARCIA-GONZALEZ, M. M., MARTINEZ-PASCUAL, M., AVELLI, P., MARTINEZ-MARTINEZ, R., FABREGA, R., COLOMER-VIROSTA, C., SORIANO-RODRIGUEZ, P., DIAZ, M., DE ZEGHER, F., IBANEZ, L. & LOPEZ-BERMEJO, A. 2010. Placental FTO expression relates to fetal growth. *Int J Obes (Lond),* 34, 1365-70.

BEAMER, B. A., NEGRI, C., YEN, C. J., GAVRILOVA, O., RUMBERGER, J. M., DURCAN, M. J., YARNALL, D. P., HAWKINS, A. L., GRIFFIN, C. A., BURNS, D. K., ROTH, J., REITMAN, M. & SHULDINER, A. R. 1997. Chromosomal localization and partial genomic structure of the human peroxisome proliferator activated receptor-gamma (hPPAR gamma) gene. *Biochem Biophys Res Commun,* 233, 756-9.

BEAULIEU, J. F., MENARD, D. & CALVERT, R. 1985. Influence of epidermal growth factor on the maturation of the fetal mouse duodenum in organ culture. *J Pediatr Gastroenterol Nutr,* 4, 476-81.

BEECHEY, C. V. 2004. A reassessment of imprinting regions and phenotypes on mouse chromosome 6: Nap1l5 locates within the currently defined sub-proximal imprinting region. *Cytogenet Genome Res,* 107, 108-14.

BELL, G. I., FONG, N. M., STEMPIEN, M. M., WORMSTED, M. A., CAPUT, D., KU, L. L., URDEA, M. S., RALL, L. B. & SANCHEZ-PESCADOR, R. 1986. Human epidermal growth factor precursor: cDNA sequence,

expression in vitro and gene organization. *Nucleic Acids Res,* 14, 8427-46.

BELLAMY, L., CASAS, J. P., HINGORANI, A. D. & WILLIAMS, D. J. 2007. Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *BMJ,* 335, 974.

BENEDICT, A. D., MARY ANN, O. R., KIRCHNER, H. L., DINESH, S. & MAUREEN, H. 2001. Perinatal correlates and neonatal outcomes of small for gestational age infants born at term gestation. *American Journal of Obstetrics and Gynecology,* 185, 652-659.

BENNETT, A. J., SOVIO, U., RUOKONEN, A., MARTIKAINEN, H., POUTA, A., HARTIKAINEN, A. L., FRANKS, S., ELLIOTT, P., JARVELIN, M. R. & MCCARTHY, M. I. 2008. No evidence that established type 2 diabetes susceptibility variants in the PPARG and KCNJ11 genes have pleiotropic effects on early growth. *Diabetologia,* 51, 82-5.

BERARDI, P., MEYYAPPAN, M. & RIABOWOL, K. T. 2003. A novel transcriptional inhibitory element differentially regulates the cyclin D1 gene in senescent cells. *J Biol Chem,* 278, 7510-9.

BERSINGER, N. A. & ODEGARD, R. A. 2004. Second- and third-trimester serum levels of placental proteins in preeclampsia and small-for-gestational age pregnancies. *Acta Obstet Gynecol Scand,* 83, 37-45.

BHATTACHARYA, S., CAMPBELL, D. M. & LISTON, W. A. 2007. Effect of Body Mass Index on pregnancy outcomes in nulliparous women delivering singleton babies. *BMC Public Health,* 7, 168.

BHOWMICK, D. A., ZHUANG, Z., WAIT, S. D. & WEIL, R. J. 2004. A functional polymorphism in the EGF gene is found with increased frequency in glioblastoma multiforme patients and is associated with more aggressive disease. *Cancer Res,* 64, 1220-3.

BIELINSKI, S. J., PANKOW, J. S., FOLSOM, A. R., NORTH, K. E. & BOERWINKLE, E. 2008. TCF7L2 single nucleotide polymorphisms, cardiovascular disease and all-cause mortality: the Atherosclerosis Risk in Communities (ARIC) study. *Diabetologia,* 51, 968-70.

BILSLAND, A. E., STEVENSON, K., ATKINSON, S., KOLCH, W. & KEITH, W. N. 2006. Transcriptional repression of telomerase RNA gene expression by c-Jun-NH2-kinase and Sp1/Sp3. *Cancer Res,* 66, 1363-70.

BISSONNETTE, F., COOK, C., GEOGHEGAN, T., STEFFEN, M., HENRY, J., YUSSMAN, M. A. & SCHULTZ, G. 1992. Transforming growth factor-alpha and epidermal growth factor messenger ribonucleic acid and protein levels in human placentas from early, mid, and late gestation. *Am J Obstet Gynecol,* 166, 192-9.

BJORCK, H. M., LANNE, T., ALEHAGEN, U., PERSSON, K., RUNDKVIST, L., HAMSTEN, A., DAHLSTROM, U. & ERIKSSON, P. 2009. Association of genetic variation on chromosome 9p21.3 and arterial stiffness. *J Intern Med,* 265, 373 - 81.

BLENCOWE, B. J. 2000. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends Biochem Sci,* 25, 106-10.

BODNAR, L. M., NESS, R. B., MARKOVIC, N. & ROBERTS, J. M. 2005. The risk of preeclampsia rises with increasing prepregnancy body mass index. *Ann Epidemiol,* 15, 475-82.

BOGDAN, S. & KLAMBT, C. 2001. Epidermal growth factor receptor signaling. *Curr Biol,* 11, R292-5.

BOISSEL, S., REISH, O., PROULX, K., KAWAGOE-TAKAKI, H., SEDGWICK, B., YEO, G. S., MEYRE, D., GOLZIO, C., MOLINARI, F., KADHOM, N., ETCHEVERS, H. C., SAUDEK, V., FAROOQI, I. S., FROGUEL, P., LINDAHL, T., O'RAHILLY, S., MUNNICH, A. & COLLEAUX, L. 2009. Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. *Am J Hum Genet,* 85, 106-11.

BONKOBARA, M., THONGSONG, B., OHMORI, T., MATSUKI, N., INABA, M. & ONO, K. 2002. Effects of epidermal growth factor on placental amino acids uptake in pregnant rats. *J Vet Med Sci,* 64, 689-92.

BONNEFOND, A., FROGUEL, P. & VAXILLAIRE, M. 2010. The emerging genetics of type 2 diabetes. *Trends Mol Med,* 16, 407-16.

BOONSTRA, J., MUMMERY, C. L., VAN DER SAAG, P. T. & DE LAAT, S. W. 1985. Two receptor classes for epidermal growth factor on pheochromocytoma cells, distinguishable by temperature, lectins, and tumor promoters. *J Cell Physiol,* 123, 347-52.

BOUHOURS-NOUET, N., MAY-PANLOUP, P., COUTANT, R., DE CASSON, F. B., DESCAMPS, P., DOUAY, O., REYNIER, P., RITZ, P., MALTHIERY, Y. & SIMARD, G. 2005. Maternal smoking is associated with mitochondrial DNA depletion and respiratory chain complex III deficiency in placenta. *Am J Physiol Endocrinol Metab,* 288, E171-7.

BOURQUE, D. K., AVILA, L., PENAHERRERA, M., VON DADELSZEN, P. & ROBINSON, W. P. 2010. Decreased placental methylation at the H19/IGF2 imprinting control region is associated with normotensive intrauterine growth restriction but not preeclampsia. *Placenta,* 31, 197-202.

BRAGT, M. C. & POPEIJUS, H. E. 2008. Peroxisome proliferator-activated receptors and the metabolic syndrome. *Physiol Behav,* 94, 187-97.

BREMME, K. A. 2003. Haemostatic changes in pregnancy. *Best Pract Res Clin Haematol,* 16, 153-68.

BROADBENT, H. M., PEDEN, J. F., LORKOWSKI, S., GOEL, A., ONGEN, H., GREEN, F., CLARKE, R., COLLINS, R., FRANZOSI, M. G., TOGNONI, G., SEEDORF, U., RUST, S., ERIKSSON, P., HAMSTEN, A., FARRALL, M., WATKINS, H. & FOR THE, P. C. 2008. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum. Mol. Genet.,* 17, 806-814.

BROUGHTON-PIPKIN, F. 2008. Smoking in moderate/severe preeclampsia worsens pregnancy outcome, but smoking cessation limits the damage. *Hypertension,* 51, 1042-6.

BROWN, M. A., LINDHEIMER, M. D., DE SWIET, M., VAN ASSCHE, A. & MOUTQUIN, J. M. 2001. The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Hypertens Pregnancy,* 20, IX-XIV.

BUCHBINDER, A., SIBAI, B. M., CARITIS, S., MACPHERSON, C., HAUTH, J., LINDHEIMER, M. D., KLEBANOFF, M., VANDORSTEN, P., LANDON, M., PAUL, R., MIODOVNIK, M., MEIS, P. & THURNAU, G. 2002. Adverse perinatal outcomes are significantly higher in severe

gestational hypertension than in mild preeclampsia. *Am J Obstet Gynecol,* 186, 66-71.

BUCHMILLER, T. L., SHAW, K. S., CHOPOURIAN, H. L., LLOYD, K. C., GREGG, J. P., RIVERA, F. A., JR., LAM, M. L., DIAMOND, J. M. & FONKALSRUD, E. W. 1993. Effect of transamniotic administration of epidermal growth factor on fetal rabbit small intestinal nutrient transport and disaccharidase development. *J Pediatr Surg,* 28, 1239-44.

BUCKLAND, P. R., HOOGENDOORN, B., COLEMAN, S. L., GUY, C. A., SMITH, S. K. & O'DONOVAN, M. C. 2005. Strong bias in the location of functional promoter polymorphisms. *Hum Mutat,* 26, 214-23.

BURD, C. E., JECK, W. R., LIU, Y., SANOFF, H. K., WANG, Z. & SHARPLESS, N. E. 2010. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet,* 6, e1001233.

BUTTE, N. F. 2000. Carbohydrate and lipid metabolism in pregnancy: normal compared with gestational diabetes mellitus. *Am J Clin Nutr,* 71, 1256S-61S.

CAI, G., COLE, S. A., HAACK, K., BUTTE, N. F. & COMUZZIE, A. G. 2007. Bivariate linkage confirms genetic contribution to fetal origins of childhood growth and cardiovascular disease risk in Hispanic children. *Hum Genet,* 121, 737-44.

CALVO, M. T., ROMO, A., GUTIERREZ, J. J., RELANO, E., BARRIO, E. & FERRANDEZ LONGAS, A. 2004. Study of genetic expression of intrauterine growth factors IGF-I and EGFR in placental tissue from pregnancies with intrauterine growth retardation. *J Pediatr Endocrinol Metab,* 17 Suppl 3, 445-50.

CARR, D. B., NEWTON, K. M., UTZSCHNEIDER, K. M., TONG, J., GERCHMAN, F., KAHN, S. E., EASTERLING, T. R. & HECKBERT, S. R. 2009. Preeclampsia and risk of developing subsequent diabetes. *Hypertens Pregnancy,* 28, 435-47.

CARTEGNI, L., CHEW, S. L. & KRAINER, A. R. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet,* 3, 285-98.

CARTEGNI, L. & KRAINER, A. R. 2002. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nat Genet,* 30, 377-84.

CARTEGNI, L., WANG, J., ZHU, Z., ZHANG, M. Q. & KRAINER, A. R. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res,* 31, 3568-71.

CASEY, M. L., KORTE, K. & MACDONALD, P. C. 1988. Epidermal growth factor stimulation of prostaglandin E2 biosynthesis in amnion cells. Induction of prostaglandin H2 synthase. *J. Biol. Chem.,* 263, 7846-7854.

CATALANO, P. M. 2010. Obesity, insulin resistance, and pregnancy outcome. *Reproduction,* 140, 365-71.

CAUCHI, S., CHOQUET, H., GUTIERREZ-AGUILAR, R., CAPEL, F., GRAU, K., PROENCA, C., DINA, C., DUVAL, A., BALKAU, B., MARRE, M., POTOCZNA, N., LANGIN, D., HORBER, F., SORENSEN, T. I. A., CHARPENTIER, G., MEYRE, D. & FROGUEL, P. 2008. Effects of TCF7L2 Polymorphisms on Obesity in European Populations. *Obesity,* 16, 476-482.

CAUCHI, S., EL ACHHAB, Y., CHOQUET, H., DINA, C., KREMPLER, F., WEITGASSER, R., NEJJARI, C., PATSCH, W., CHIKRI, M., MEYRE, D. & FROGUEL, P. 2007a. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med,* 85, 777-82.

CAUCHI, S., MEYRE, D., CHOQUET, H., DEGHMOUN, S., DURAND, E., GAGET, S., LECOEUR, C., FROGUEL, P. & LEVY-MARCHAL, C. 2007b. TCF7L2 rs7903146 variant does not associate with smallness for gestational age in the French population. *BMC Med Genet,* 8, 37.

CAUCHI, S., MEYRE, D., DINA, C., CHOQUET, H., SAMSON, C., GALLINA, S., BALKAU, B., CHARPENTIER, G., PATTOU, F., STETSYUK, V., SCHARFMANN, R., STAELS, B., FRUHBECK, G. & FROGUEL, P. 2006. Transcription Factor TCF7L2 Genetic Study in the French Population: Expression in Human {beta}-Cells and Adipose Tissue and Strong Association With Type 2 Diabetes. *Diabetes,* 55, 2903-2908.

CAZZOLA, M. & SKODA, R. C. 2000. Translational pathophysiology: a novel molecular mechanism of human disease. *Blood,* 95, 3280-8.

CECIL, J. E., TAVENDALE, R., WATT, P., HETHERINGTON, M. M. & PALMER, C. N. 2008. An obesity-associated FTO gene variant and increased energy intake in children. *N Engl J Med,* 359, 2558-66.

CELLINI, C., XU, J., ARRIAGA, A. & BUCHMILLER-CRAIR, T. L. 2004. Effect of epidermal growth factor infusion on fetal rabbit intrauterine growth retardation and small intestinal development. *J Pediatr Surg,* 39, 891-7; discussion 891-7.

CETIN, I., FOIDART, J. M., MIOZZO, M., RAUN, T., JANSSON, T., TSATSARIS, V., REIK, W., CROSS, J., HAUGUEL-DE-MOUZON, S., ILLSLEY, N., KINGDOM, J. & HUPPERTZ, B. 2004. Fetal growth restriction: a workshop report. *Placenta,* 25, 753-757.

CETIN, I., GIOVANNINI, N., ALVINO, G., AGOSTONI, C., RIVA, E., GIOVANNINI, M. & PARDI, G. 2002. Intrauterine growth restriction is associated with changes in polyunsaturated fatty acid fetal-maternal relationships. *Pediatr Res,* 52, 750-5.

CHAJRY, N., MARTIN, P. M., PAGES, G., COCHET, C., AFDEL, K. & BERTHOIS, Y. 1994. Relationship between the MAP kinase activity and the dual effect of EGF on A431 cell proliferation. *Biochem Biophys Res Commun,* 203, 984-90.

CHAKRAVARTY, E. F., KHANNA, D. & CHUNG, L. 2008. Pregnancy outcomes in systemic sclerosis, primary pulmonary hypertension, and sickle cell disease. *Obstet Gynecol,* 111, 927-34.

CHAN, S. Y. & WONG, R. W. 2000. Expression of epidermal growth factor in transgenic mice causes growth retardation. *J Biol Chem,* 275, 38693-8.

CHAPMAN, J. M., COOPER, J. D., TODD, J. A. & CLAYTON, D. G. 2003. Detecting disease associations due to linkage disequilibrium using haplotype tags: a class of tests and the determinants of statistical power. *Hum Hered,* 56, 18-31.

CHAPPELL, L. C., ENYE, S., SEED, P., BRILEY, A. L., POSTON, L. & SHENNAN, A. H. 2008. Adverse Perinatal Outcomes and Risk Factors for Preeclampsia in Women With Chronic Hypertension: A Prospective Study. *Hypertension,* 51, 1002-1009.

CHAPPELL, S. & MORGAN, L. 2006. Searching for genetic clues to the causes of pre-eclampsia. *Clin Sci (Lond),* 110, 443-58.

CHATTERJEE, S. & PAL, J. K. 2009. Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biol Cell,* 101, 251-62.

CHEN, Q., DING, J. X., LIU, B., STONE, P., FENG, Y. J. & CHAMLEY, L. 2010. Spreading endothelial cell dysfunction in response to necrotic trophoblasts. Soluble factors released from endothelial cells that have phagocytosed necrotic shed trophoblasts reduce the proliferation of additional endothelial cells. *Placenta,* 31, 976-81.

CHEN, Y., JIMENEZ, A. R. & MEDH, J. D. 2006. Identification and regulation of novel PPAR-gamma splice variants in human THP-1 macrophages. *Biochim Biophys Acta,* 1759, 32-43.

CHISTIAKOV, D. A., POTAPOV, V. A., KHODIREV, D. C., SHAMKHALOVA, M. S., SHESTAKOVA, M. V. & NOSIKOV, V. V. 2009. Genetic variations in the pancreatic ATP-sensitive potassium channel, beta-cell dysfunction, and susceptibility to type 2 diabetes. *Acta Diabetol,* 46, 43-9.

CHO, Y. M., KIM, T. H., LIM, S., CHOI, S. H., SHIN, H. D., LEE, H. K., PARK, K. S. & JANG, H. C. 2009. Type 2 diabetes-associated genetic variants discovered in the recent genome-wide association studies are related to gestational diabetes mellitus in the Korean population. *Diabetologia,* 52, 253-61.

CHURCH, C., LEE, S., BAGG, E. A., MCTAGGART, J. S., DEACON, R., GERKEN, T., LEE, A., MOIR, L., MECINOVIC, J., QUWAILID, M. M., SCHOFIELD, C. J., ASHCROFT, F. M. & COX, R. D. 2009. A mouse model for the metabolic effects of the human fat mass and obesity associated FTO gene. *PLoS Genet,* 5, e1000599.

CIRULLI, E. T. & GOLDSTEIN, D. B. 2007. In vitro assays fail to predict in vivo effects of regulatory polymorphisms. *Hum Mol Genet,* 16, 1931-9.

CLAUSEN, T., DJUROVIC, S. & HENRIKSEN, T. 2001. Dyslipidemia in early second trimester is mainly a feature of women with early onset pre-eclampsia. *BJOG,* 108, 1081-7.

CNATTINGIUS, S., REILLY, M., PAWITAN, Y. & LICHTENSTEIN, P. 2004. Maternal and fetal genetic factors account for most of familial aggregation of preeclampsia: a population-based Swedish cohort study. *Am J Med Genet A,* 130A, 365-71.

COOPER-DEHOFF, R. M., PEPINE, C. J. 2007. Metabolic Syndrome and Cardiovascular Disease: Challenges and Opportunities. *Clin. Cardiol.* , 30, 593–597

COOPER, D. W. & LISTON, W. A. 1979. Genetic control of severe pre-eclampsia. *J Med Genet,* 16, 409-16.

COOPER, T. A. 2005. Use of minigene systems to dissect alternative splicing elements. *Methods,* 37, 331-340.

CORDELL, H. J. & CLAYTON, D. G. 2005. Genetic association studies. *Lancet,* 366, 1121-31.

COSTA, B. M., FERREIRA, P., COSTA, S., CANEDO, P., OLIVEIRA, P., SILVA, A., PARDAL, F., SURIANO, G., MACHADO, J. C., LOPES, J. M. & REIS, R. M. 2007. Association between functional EGF+61 polymorphism and glioma risk. *Clin Cancer Res,* 13, 2621-6.

COTE, A. M., FIROZ, T., MATTMAN, A., LAM, E. M., VON DADELSZEN, P. & MAGEE, L. A. 2008. The 24-hour urine collection: gold standard or historical practice? *Am J Obstet Gynecol,* 199, 625 e1-6.

CRAICI, I., WAGNER, S. & GAROVIC, V. D. 2008. Preeclampsia and future cardiovascular risk: formal risk factor or failed stress test? *Ther Adv Cardiovasc Dis,* 2, 249-59.

CROCKER, I. P., COOPER, S., ONG, S. C. & BAKER, P. N. 2003. Differences in apoptotic susceptibility of cytotrophoblasts and syncytiotrophoblasts in normal pregnancy to those complicated with preeclampsia and intrauterine growth restriction. *Am J Pathol,* 162, 637-43.

CUNNINGTON, M. S., SANTIBANEZ KOREF, M., MAYOSI, B. M., BURN, J. & KEAVNEY, B. 2010. Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. *PLoS Genet,* 6, e1000899.

DASH, P. R., WHITLEY, G. S., AYLING, L. J., JOHNSTONE, A. P. & CARTWRIGHT, J. E. 2005. Trophoblast apoptosis is inhibited by hepatocyte growth factor through the Akt and beta-catenin mediated up-regulation of inducible nitric oxide synthase. *Cell Signal,* 17, 571-80.

DAWN TEARE, M. & BARRETT, J. H. 2005. Genetic linkage studies. *Lancet,* 366, 1036-44.

DEEB, S. S., FAJAS, L., NEMOTO, M., PIHLAJAMAKI, J., MYKKANEN, L., KUUSISTO, J., LAAKSO, M., FUJIMOTO, W. & AUWERX, J. 1998. A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. *Nat Genet,* 20, 284-7.

DEFRONZO, R. A., BONADONNA, R. C. & FERRANNINI, E. 1992. Pathogenesis of NIDDM. A balanced overview. *Diabetes Care,* 15, 318-68.

DEJMEK, J., SOLANSK Y, I., PODRAZILOVA, K. & SRAM, R. J. 2002. The exposure of nonsmoking and smoking mothers to environmental tobacco smoke during different gestational phases and fetal growth. *Environ Health Perspect,* 110, 601-6.

DEPALO, L. & DAS, M. 1988. Epidermal growth factor-induced stimulation of epidermal growth factor-receptor synthesis in human cytotrophoblasts and A431 carcinoma cells. *Cancer Res,* 48, 1105-9.

DEWEY, C. N., ROGOZIN, I. B. & KOONIN, E. V. 2006. Compensatory relationship between splice sites and exonic splicing signals depending on the length of vertebrate introns. *BMC Genomics,* 7, 311.

DINA, C., MEYRE, D., GALLINA, S., DURAND, E., KORNER, A., JACOBSON, P., CARLSSON, L. M., KIESS, W., VATIN, V., LECOEUR, C., DELPLANQUE, J., VAILLANT, E., PATTOU, F., RUIZ, J., WEILL, J., LEVY-MARCHAL, C., HORBER, F., POTOCZNA, N., HERCBERG, S., LE STUNFF, C., BOUGNERES, P., KOVACS, P., MARRE, M., BALKAU, B., CAUCHI, S., CHEVRE, J. C. & FROGUEL, P. 2007. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet,* 39, 724-6.

DISSANAYAKE, V. H., GILES, V., JAYASEKARA, R. W., SENEVIRATNE, H. R., KALSHEKER, N., BROUGHTON PIPKIN, F. & MORGAN, L. 2009. A study of three candidate genes for pre-eclampsia in a Sinhalese population from Sri Lanka. *J Obstet Gynaecol Res,* 35, 234-42.

DISSANAYAKE, V. H., TOWER, C., BRODERICK, A., STOCKER, L. J., SENEVIRATNE, H. R., JAYASEKARA, R. W., KALSHEKER, N., BROUGHTON PIPKIN, F. & MORGAN, L. 2007. Polymorphism in the epidermal growth factor gene is associated with birthweight in Sinhalese and white Western Europeans. *Mol Hum Reprod,* 13, 425-9.

DISSET, A., BOURGEOIS, C. F., BENMALEK, N., CLAUSTRES, M., STEVENIN, J. & TUFFERY-GIRAUD, S. 2006. An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements. *Hum Mol Genet,* 15, 999-1013.

DO, R., BAILEY, S. D., DESBIENS, K., BELISLE, A., MONTPETIT, A., BOUCHARD, C., PERUSSE, L., VOHL, M.-C. & ENGERT, J. C. 2008. Genetic Variants of FTO Influence Adiposity, Insulin Sensitivity, Leptin Levels, and Resting Metabolic Rate in the Quebec Family Study. *Diabetes,* 57, 1147-1150.

DOI, Y., KUBO, M., NINOMIYA, T., YONEMOTO, K., IWASE, M., ARIMA, H., HATA, J., TANIZAKI, Y., IIDA, M. & KIYOHARA, Y. 2007. Impact of Kir6.2 E23K polymorphism on the development of type 2 diabetes in a general Japanese population: The Hisayama Study. *Diabetes,* 56, 2829-33.

DONEY, A., FISCHER, B., FREW, D., CUMMING, A., FLAVELL, D. M., WORLD, M., MONTGOMERY, H. E., BOYLE, D., MORRIS, A. & PALMER, C. N. 2002. Haplotype analysis of the PPARgamma Pro12Ala and C1431T variants reveals opposing associations with body weight. *BMC Genet,* 3, 21.

DONEY, A. S., DANNFALD, J., KIMBER, C. H., DONNELLY, L. A., PEARSON, E., MORRIS, A. D. & PALMER, C. N. 2009. The FTO gene is associated with an atherogenic lipid profile and myocardial infarction in patients with type 2 diabetes: a Genetics of Diabetes Audit and Research Study in Tayside Scotland (Go-DARTS) study. *Circ Cardiovasc Genet,* 2, 255-9.

DORIA, A., WOJCIK, J., XU, R., GERVINO, E. V., HAUSER, T. H., JOHNSTONE, M. T., NOLAN, D., HU, F. B. & WARRAM, J. H. 2008. Interaction between poor glycemic control and 9p21 locus on risk of coronary artery disease in type 2 diabetes. *JAMA,* 300, 2389 - 97.

DUCKITT, K. & HARRINGTON, D. 2005. Risk factors for pre-eclampsia at antenatal booking: systematic review of controlled studies. *BMJ,* 330, 565.

DUELLO, T. M., BERTICS, P. J., FULGHAM, D. L. & VAN ESS, P. J. 1994. Localization of epidermal growth factor receptors in first- and third-trimester human placentas. *J Histochem Cytochem,* 42, 907-15.

DUESING, K., FATEMIFAR, G., CHARPENTIER, G., MARRE, M., TICHET, J., HERCBERG, S., BALKAU, B., FROGUEL, P. & GIBSON, F. 2008. Strong association of common variants in the CDKN2A/CDKN2B region with type 2 diabetes in French Europids. *Diabetologia,* 51, 821-6.

DULEY, L. 2009. The global impact of pre-eclampsia and eclampsia. *Semin Perinatol,* 33, 130-7.

DUNNE, F., BRYDON, P., SMITH, K. & GEE, H. 2003. Pregnancy in women with Type 2 diabetes: 12 years outcome data 1990-2002. *Diabet Med,* 20, 734-8.

DUVAL, A., BUSSON-LECONIAT, M., BERGER, R. & HAMELIN, R. 2000. Assignment of the TCF-4 gene (TCF7L2) to human chromosome band 10q25.3. *Cytogenet Cell Genet,* 88, 264-5.

EBNER, R. & DERYNCK, R. 1991. Epidermal growth factor and transforming growth factor-alpha: differential intracellular routing and processing of ligand-receptor complexes. *Cell Regul,* 2, 599-612.

EGBOR, M., ANSARI, T., MORRIS, N., GREEN, C. J. & SIBBONS, P. D. 2006a. Morphometric placental villous and vascular abnormalities in early- and late-onset pre-eclampsia with and without fetal growth restriction. *BJOG,* 113, 580-9.

EGBOR, M., ANSARI, T., MORRIS, N., GREEN, C. J. & SIBBONS, P. D. 2006b. Pre-eclampsia and fetal growth restriction: how morphometrically different is the placenta? *Placenta,* 27, 727-34.

EK, J., ANDERSEN, G., URHAMMER, S. A., HANSEN, L., CARSTENSEN, B., BORCH-JOHNSEN, K., DRIVSHOLM, T., BERGLUND, L., HANSEN, T., LITHELL, H. & PEDERSEN, O. 2001. Studies of the Pro12Ala polymorphism of the peroxisome proliferator-activated receptor-gamma2 (PPAR-gamma2) gene in relation to insulin sensitivity among glucose tolerant caucasians. *Diabetologia,* 44, 1170-6.

EK, J., URHAMMER, S. A., SORENSEN, T. I., ANDERSEN, T., AUWERX, J. & PEDERSEN, O. 1999. Homozygosity of the Pro12Ala variant of the peroxisome proliferation-activated receptor-gamma2 (PPAR-gamma2): divergent modulating effects on body mass index in obese and lean Caucasian men. *Diabetologia,* 42, 892-5.

ELBRECHT, A., CHEN, Y., CULLINAN, C. A., HAYES, N., LEIBOWITZ, M., MOLLER, D. E. & BERGER, J. 1996. Molecular cloning, expression and characterization of human peroxisome proliferator activated receptors gamma 1 and gamma 2. *Biochem Biophys Res Commun,* 224, 431-7.

EMERY, S. P., LEVINE, R. J., QIAN, C., EWELL, M. G., ENGLAND, L. J., YU, K. F. & CATALANO, P. M. 2005. Twenty-four-hour urine insulin as a measure of hyperinsulinaemia/insulin resistance before onset of pre-eclampsia and gestational hypertension. *BJOG,* 112, 1479-85.

ENGLAND, L. J., LEVINE, R. J., QIAN, C., MORRIS, C. D., SIBAI, B. M., CATALANO, P. M., CURET, L. B. & KLEBANOFF, M. A. 2002. Smoking before pregnancy and risk of gestational hypertension and preeclampsia. *Am J Obstet Gynecol,* 186, 1035-40.

EREZ, O., ROMERO, R., ESPINOZA, J., FU, W., TODEM, D., KUSANOVIC, J. P., GOTSCH, F., EDWIN, S., NIEN, J. K., CHAIWORAPONGSA, T., MITTAL, P., MAZAKI-TOVI, S., THAN, N. G., GOMEZ, R. & HASSAN, S. S. 2008. The change in concentrations of angiogenic and anti-angiogenic factors in maternal plasma between the first and second trimesters in risk assessment for the subsequent development of preeclampsia and small-for-gestational age. *J Matern Fetal Neonatal Med,* 21, 279-87.

ERIKSSON, J. G. 2007. Epidemiology, genes and the environment: lessons learned from the Helsinki Birth Cohort Study. *Journal of Internal Medicine,* 261, 418-425.

ERIKSSON, J. G., LINDI, V., UUSITUPA, M., FORSEN, T. J., LAAKSO, M., OSMOND, C. & BARKER, D. J. 2002. The effects of the Pro12Ala polymorphism of the peroxisome proliferator-activated receptor-gamma2

gene on insulin sensitivity and insulin metabolism interact with size at birth. *Diabetes,* 51, 2321-4.

ESCHER, P. & WAHLI, W. 2000. Peroxisome proliferator-activated receptors: insight into multiple cellular functions. *Mutat Res,* 448, 121-38.

ESLICK, G. D., YAN, P., XIA, H. H., MURRAY, H., SPURRETT, B. & TALLEY, N. J. 2002. Foetal intrauterine growth restrictions with Helicobacter pylori infection. *Aliment Pharmacol Ther,* 16, 1677-82.

ESPLIN, M. S., FAUSETT, M. B., FRASER, A., KERBER, R., MINEAU, G., CARRILLO, J. & VARNER, M. W. 2001. Paternal and maternal components of the predisposition to preeclampsia. *N Engl J Med,* 344, 867-72.

EXPERT-BEZANCON, A., SUREAU, A., DUROSAY, P., SALESSE, R., GROENEVELD, H., LECAER, J. P. & MARIE, J. 2004. hnRNP A1 and the SR proteins ASF/SF2 and SC35 have antagonistic functions in splicing of beta-tropomyosin exon 6B. *J Biol Chem,* 279, 38249-59.

FABIO GHEZZI, M. G. T. L. R. E. D. N. B. L. M. T. M. F. 2003. Idiopathic fetal intrauterine growth restriction: a possible inheritance pattern. *Prenatal Diagnosis,* 23, 259-264.

FAIRBROTHER, W. G., YEH, R. F., SHARP, P. A. & BURGE, C. B. 2002. Predictive identification of exonic splicing enhancers in human genes. *Science,* 297, 1007-13.

FAIRBROTHER, W. G., YEO, G. W., YEH, R., GOLDSTEIN, P., MAWSON, M., SHARP, P. A. & BURGE, C. B. 2004. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res,* 32, W187-90.

FERRANDINA, G., LANZONE, A., SCAMBIA, G., CARUSO, A., PANICI, P. B. & MANCUSO, S. 1995. Epidermal growth factor receptors in placentae and fetal membranes from hypertension-complicated pregnancies. *Hum Reprod,* 10, 1845-9.

FISCHER, J., KOCH, L., EMMERLING, C., VIERKOTTEN, J., PETERS, T., BRUNING, J. C. & RUTHER, U. 2009. Inactivation of the Fto gene protects from obesity. *Nature,* 458, 894-8.

FISHER, D. A. & LAKSHMANAN, J. 1990. Metabolism and effects of epidermal growth factor and related growth factors in mammals. *Endocr Rev,* 11, 418-442.

FITZPATRICK, E., JOHNSON, M. P., DYER, T. D., FORREST, S., ELLIOTT, K., BLANGERO, J., BRENNECKE, S. P. & MOSES, E. K. 2009. Genetic association of the activin A receptor gene (ACVR2A) and pre-eclampsia. *Mol Hum Reprod,* 15, 195-204.

FLOREZ, J. C., BURTT, N., DE BAKKER, P. I., ALMGREN, P., TUOMI, T., HOLMKVIST, J., GAUDET, D., HUDSON, T. J., SCHAFFNER, S. F., DALY, M. J., HIRSCHHORN, J. N., GROOP, L. & ALTSHULER, D. 2004. Haplotype structure and genotype-phenotype correlations of the sulfonylurea receptor and the islet ATP-sensitive potassium channel gene region. *Diabetes,* 53, 1360-8.

FLOREZ, J. C., JABLONSKI, K. A., KAHN, S. E., FRANKS, P. W., DABELEA, D., HAMMAN, R. F., KNOWLER, W. C., NATHAN, D. M. & ALTSHULER, D. 2007. Type 2 diabetes-associated missense polymorphisms KCNJ11 E23K and ABCC8 A1369S influence progression to diabetes and response to interventions in the Diabetes Prevention Program. *Diabetes,* 56, 531-6.

FLORIAN, J. A. & WATTS, S. W. 1999. Epidermal growth factor: a potent vasoconstrictor in experimental hypertension. *American Journal of Physiology - Heart and Circulatory Physiology,* 276, H976-H983.

FOLGERO, T., STORBAKK, N., TORBERGSEN, T. & OIAN, P. 1996. Mutations in mitochondrial transfer ribonucleic acid genes in preeclampsia. *Am J Obstet Gynecol,* 174, 1626-30.

FONDACCI, C., ALSAT, E., GABRIEL, R., BLOT, P., NESSMANN, C. & EVAIN-BRION, D. 1994. Alterations of human placental epidermal growth factor receptor in intrauterine growth retardation. *J Clin Invest,* 93, 1149-55.

FRAYLING, T. M. & HATTERSLEY, A. T. 2001. The role of genetic susceptibility in the association of low birth weight with type 2 diabetes.

FRAYLING, T. M., TIMPSON, N. J., WEEDON, M. N., ZEGGINI, E., FREATHY, R. M., LINDGREN, C. M., PERRY, J. R. B., ELLIOTT, K. S., LANGO, H., RAYNER, N. W., SHIELDS, B., HARRIES, L. W., BARRETT, J. C., ELLARD, S., GROVES, C. J., KNIGHT, B., PATCH, A.-M., NESS, A. R., EBRAHIM, S., LAWLOR, D. A., RING, S. M., BEN-SHLOMO, Y., JARVELIN, M.-R., SOVIO, U., BENNETT, A. J., MELZER, D., FERRUCCI, L., LOOS, R. J. F., BARROSO, I., WAREHAM, N. J., KARPE, F., OWEN, K. R., CARDON, L. R., WALKER, M., HITMAN, G. A., PALMER, C. N. A., DONEY, A. S. F., MORRIS, A. D., SMITH, G. D., THE WELLCOME TRUST CASE CONTROL, C., HATTERSLEY, A. T. & MCCARTHY, M. I. 2007. A Common Variant in the FTO Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. *Science,* 316, 889-894.

FREATHY, R. M., BENNETT, A. J., RING, S. M., SHIELDS, B., GROVES, C. J., TIMPSON, N. J., WEEDON, M. N., ZEGGINI, E., LINDGREN, C. M., LANGO, H., PERRY, J. R., POUTA, A., RUOKONEN, A., HYPPONEN, E., POWER, C., ELLIOTT, P., STRACHAN, D. P., JARVELIN, M. R., SMITH, G. D., MCCARTHY, M. I., FRAYLING, T. M. & HATTERSLEY, A. T. 2009. Type 2 diabetes risk alleles are associated with reduced size at birth. *Diabetes,* 58, 1428-33.

FREATHY, R. M., MOOK-KANAMORI, D. O., SOVIO, U., PROKOPENKO, I., TIMPSON, N. J., BERRY, D. J., WARRINGTON, N. M., WIDEN, E., HOTTENGA, J. J., KAAKINEN, M., LANGE, L. A., BRADFIELD, J. P., KERKHOF, M., MARSH, J. A., MAGI, R., CHEN, C. M., LYON, H. N., KIRIN, M., ADAIR, L. S., AULCHENKO, Y. S., BENNETT, A. J., BORJA, J. B., BOUATIA-NAJI, N., CHAROEN, P., COIN, L. J., COUSMINER, D. L., DE GEUS, E. J., DELOUKAS, P., ELLIOTT, P., EVANS, D. M., FROGUEL, P., GLASER, B., GROVES, C. J., HARTIKAINEN, A. L., HASSANALI, N., HIRSCHHORN, J. N., HOFMAN, A., HOLLY, J. M., HYPPONEN, E., KANONI, S., KNIGHT, B. A., LAITINEN, J., LINDGREN, C. M., MCARDLE, W. L., O'REILLY, P. F., PENNELL, C. E., POSTMA, D. S., POUTA, A., RAMASAMY, A., RAYNER, N. W., RING, S. M., RIVADENEIRA, F., SHIELDS, B. M., STRACHAN, D. P., SURAKKA, I., TAANILA, A., TIESLER, C., UITTERLINDEN, A. G., VAN DUIJN, C. M., WIJGA, A. H., WILLEMSEN, G., ZHANG, H., ZHAO, J., WILSON, J. F., STEEGERS, E. A., HATTERSLEY, A. T., ERIKSSON, J. G., PELTONEN, L., MOHLKE, K. L., GRANT, S. F., HAKONARSON, H., KOPPELMAN, G. H., DEDOUSSIS, G. V., HEINRICH, J., GILLMAN,

M. W., PALMER, L. J., FRAYLING, T. M., BOOMSMA, D. I., DAVEY SMITH, G., POWER, C., JADDOE, V. W., JARVELIN, M. R. & MCCARTHY, M. I. 2010. Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. *Nat Genet,* 42, 430-5.

FREATHY, R. M., TIMPSON, N. J., LAWLOR, D. A., POUTA, A., BEN-SHLOMO, Y., RUOKONEN, A., EBRAHIM, S., SHIELDS, B., ZEGGINI, E., WEEDON, M. N., LINDGREN, C. M., LANGO, H., MELZER, D., FERRUCCI, L., PAOLISSO, G., NEVILLE, M. J., KARPE, F., PALMER, C. N., MORRIS, A. D., ELLIOTT, P., JARVELIN, M. R., SMITH, G. D., MCCARTHY, M. I., HATTERSLEY, A. T. & FRAYLING, T. M. 2008. Common variation in the FTO gene alters diabetes-related metabolic traits to the extent expected given its effect on BMI. *Diabetes,* 57, 1419-26.

FREATHY, R. M., WEEDON, M. N., BENNETT, A., HYPPONEN, E., RELTON, C. L., KNIGHT, B., SHIELDS, B., PARNELL, K. S., GROVES, C. J., RING, S. M., PEMBREY, M. E., BEN-SHLOMO, Y., STRACHAN, D. P., POWER, C., JARVELIN, M. R., MCCARTHY, M. I., DAVEY SMITH, G., HATTERSLEY, A. T. & FRAYLING, T. M. 2007. Type 2 diabetes TCF7L2 risk genotypes alter birth weight: a study of 24,053 individuals. *Am J Hum Genet,* 80, 1150-61.

FREEDMAN, M. L., REICH, D., PENNEY, K. L., MCDONALD, G. J., MIGNAULT, A. A., PATTERSON, N., GABRIEL, S. B., TOPOL, E. J., SMOLLER, J. W., PATO, C. N., PATO, M. T., PETRYSHEN, T. L., KOLONEL, L. N., LANDER, E. S., SKLAR, P., HENDERSON, B., HIRSCHHORN, J. N. & ALTSHULER, D. 2004. Assessing the impact of population stratification on genetic association studies. *Nat Genet,* 36, 388-93.

FU, M., ZHANG, J., ZHU, X., MYLES, D. E., WILLSON, T. M., LIU, X. & CHEN, Y. E. 2001. Peroxisome proliferator-activated receptor gamma inhibits transforming growth factor beta-induced connective tissue growth factor expression in human aortic smooth muscle cells by interfering with Smad3. *J Biol Chem,* 276, 45888-94.

FUJITA, Y., KURACHI, H., MORISHIGE, K., AMEMIYA, K., TERAKAWA, N., MIYAKE, A. & TANIZAWA, O. 1991. Decrease in epidermal growth factor receptor and its messenger ribonucleic acid levels in intrauterine growth-retarded and diabetes mellitus-complicated pregnancies. *J Clin Endocrinol Metab,* 72, 1340-5.

GABRIEL, S. B., SCHAFFNER, S. F., NGUYEN, H., MOORE, J. M., ROY, J., BLUMENSTIEL, B., HIGGINS, J., DEFELICE, M., LOCHNER, A., FAGGART, M., LIU-CORDERO, S. N., ROTIMI, C., ADEYEMO, A., COOPER, R., WARD, R., LANDER, E. S., DALY, M. J. & ALTSHULER, D. 2002. The structure of haplotype blocks in the human genome. *Science,* 296, 2225-9.

GARDOSI, J. 2004. Customized fetal growth standards: rationale and clinical application. *Semin Perinatol,* 28, 33-40.

GARDOSI, J., CLAUSSON, B. & FRANCIS, A. 2009. The value of customised centiles in assessing perinatal mortality risk associated with parity and maternal size. *BJOG,* 116, 1356-63.

GARDOSI, J., MONGELLI, M., WILCOX, M. & CHANG, A. 1995. An adjustable fetal weight standard. *Ultrasound Obstet Gynecol,* 6, 168-74.

GARGIULO, A. R., KHAN-DAWOOD, F. S. & DAWOOD, M. Y. 1997. Epidermal growth factor receptors in uteroplacental tissues in term pregnancy before and after the onset of labor. *J Clin Endocrinol Metab,* 82, 113-7.

GAUSTER, M., MOSER, G., ORENDI, K. & HUPPERTZ, B. 2009. Factors involved in regulating trophoblast fusion: potential role in the development of preeclampsia. *Placenta,* 30 Suppl A, S49-54.

GERKEN, T., GIRARD, C. A., TUNG, Y. C., WEBBY, C. J., SAUDEK, V., HEWITSON, K. S., YEO, G. S., MCDONOUGH, M. A., CUNLIFFE, S., MCNEILL, L. A., GALVANOVSKIS, J., RORSMAN, P., ROBINS, P., PRIEUR, X., COLL, A. P., MA, M., JOVANOVIC, Z., FAROOQI, I. S., SEDGWICK, B., BARROSO, I., LINDAHL, T., PONTING, C. P., ASHCROFT, F. M., O'RAHILLY, S. & SCHOFIELD, C. J. 2007. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science,* 318, 1469-72.

GHOUSSAINI, M., MEYRE, D., LOBBENS, S., CHARPENTIER, G., CLEMENT, K., CHARLES, M.-A., TAUBER, M., WEILL, J. & FROGUEL, P. 2005. Implication of the Pro12Ala polymorphism of the PPAR-gamma 2 gene in type 2 diabetes and obesity in the French population. *BMC Medical Genetics,* 6, 11.

GIAPROS, V., DROUGIA, A., KRALLIS, N., THEOCHARIS, P. & ANDRONIKOU, S. 2011. Morbidity and mortality patterns in small-for-gestational age infants born preterm. *J Matern Fetal Neonatal Med.*

GIMELBRANT, A., HUTCHINSON, J. N., THOMPSON, B. R. & CHESS, A. 2007. Widespread monoallelic expression on human autosomes. *Science,* 318, 1136-40.

GIROUARD, J., GIGUERE, Y., MOUTQUIN, J.-M. & FOREST, J.-C. 2007. Previous Hypertensive Disease of Pregnancy Is Associated With Alterations of Markers of Insulin Resistance. *Hypertension,* 49, 1056-1062.

GLOYN, A. L., HASHIM, Y., ASHCROFT, S. J., ASHFIELD, R., WILTSHIRE, S. & TURNER, R. C. 2001. Association studies of variants in promoter and coding regions of beta-cell ATP-sensitive K-channel genes SUR1 and Kir6.2 with Type 2 diabetes mellitus (UKPDS 53). *Diabet Med,* 18, 206-12.

GLOYN, A. L., PEARSON, E. R., ANTCLIFF, J. F., PROKS, P., BRUINING, G. J., SLINGERLAND, A. S., HOWARD, N., SRINIVASAN, S., SILVA, J. M., MOLNES, J., EDGHILL, E. L., FRAYLING, T. M., TEMPLE, I. K., MACKAY, D., SHIELD, J. P., SUMNIK, Z., VAN RHIJN, A., WALES, J. K., CLARK, P., GORMAN, S., AISENBERG, J., ELLARD, S., NJOLSTAD, P. R., ASHCROFT, F. M. & HATTERSLEY, A. T. 2004. Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. *N Engl J Med,* 350, 1838-49.

GLOYN, A. L., WEEDON, M. N., OWEN, K. R., TURNER, M. J., KNIGHT, B. A., HITMAN, G., WALKER, M., LEVY, J. C., SAMPSON, M., HALFORD, S., MCCARTHY, M. I., HATTERSLEY, A. T. & FRAYLING, T. M. 2003. Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes,* 52, 568-72.

GODDARD, K. A., TROMP, G., ROMERO, R., OLSON, J. M., LU, Q., XU, Z., PARIMI, N., NIEN, J. K., GOMEZ, R., BEHNKE, E., SOLARI, M., ESPINOZA, J., SANTOLAYA, J., CHAIWORAPONGSA, T., LENK, G. M., VOLKENANT, K., ANANT, M. K., SALISBURY, B. A., CARR, J., LEE, M. S., VOVIS, G. F. & KUIVANIEMI, H. 2007. Candidate-gene association study of mothers with pre-eclampsia, and their infants, analyzing 775 SNPs in 190 genes. *Hum Hered,* 63, 1-16.

GOLDENBERG, R. L. & CLIVER, S. P. 1997. Small for gestational age and intrauterine growth restriction: definitions and standards. *Clin Obstet Gynecol,* 40, 704-14.

GONZALEZ-NAVARRO, H., ABU NABAH, Y. N., VINUE, A., ANDRES-MANZANO, M. J., COLLADO, M., SERRANO, M. & ANDRES, V. 2010. p19(ARF) deficiency reduces macrophage and vascular smooth muscle cell apoptosis and aggravates atherosclerosis. *J Am Coll Cardiol,* 55, 2258-68.

GORI, F., SPECCHIA, C., PIETRI, S., CROCIATI, L., BARLERA, S., FRANCIOSI, M., NICOLUCCI, A., SIGNORINI, S., BRAMBILLA, P., FRANZOSI, M., INVESTIGATORS, G. P. & GROUP, S.-G. P. 2010. Common genetic variants on chromosome 9p21 are associated with myocardial infarction and type 2 diabetes in an Italian population. *BMC Medical Genetics,* 11, 60.

GOSWAMI, D., TANNETTA, D. S., MAGEE, L. A., FUCHISAWA, A., REDMAN, C. W., SARGENT, I. L. & VON DADELSZEN, P. 2006. Excess syncytiotrophoblast microparticle shedding is a feature of early-onset pre-eclampsia, but not normotensive intrauterine growth restriction. *Placenta,* 27, 56-61.

GRANGER, J. P., ALEXANDER, B. T., LLINAS, M. T., BENNETT, W. A. & KHALIL, R. A. 2001. Pathophysiology of Hypertension During Preeclampsia Linking Placental Ischemia With Endothelial Dysfunction. *Hypertension,* 38, 718-722.

GRANT, S. F., THORLEIFSSON, G., REYNISDOTTIR, I., BENEDIKTSSON, R., MANOLESCU, A., SAINZ, J., HELGASON, A., STEFANSSON, H., EMILSSON, V., HELGADOTTIR, A., STYRKARSDOTTIR, U., MAGNUSSON, K. P., WALTERS, G. B., PALSDOTTIR, E., JONSDOTTIR, T., GUDMUNDSDOTTIR, T., GYLFASON, A., SAEMUNDSDOTTIR, J., WILENSKY, R. L., REILLY, M. P., RADER, D. J., BAGGER, Y., CHRISTIANSEN, C., GUDNASON, V., SIGURDSSON, G., THORSTEINSDOTTIR, U., GULCHER, J. R., KONG, A. & STEFANSSON, K. 2006. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet,* 38, 320-3.

GRIFFITHS, L. J., DEZATEUX, C. & COLE, T. J. 2007. Differential parental weight and height contributions to offspring birthweight and weight gain in infancy. *Int J Epidemiol,* 36, 104-7.

GRISSA, O., YESSOUFOU, A., MRISAK, I., HICHAMI, A., AMOUSSOU-GUENOU, D., GRISSA, A., DJROLO, F., MOUTAIROU, K., MILED, A., KHAIRI, H., ZAOUALI, M., BOUGMIZA, I., ZBIDI, A., TABKA, Z. & KHAN, N. A. 2010. Growth factor concentrations and their placental mRNA expression are modulated in gestational diabetes mellitus: possible interactions with macrosomia. *BMC Pregnancy Childbirth,* 10, 7.

GROVES, C. J., ZEGGINI, E., MINTON, J., FRAYLING, T. M., WEEDON, M. N., RAYNER, N. W., HITMAN, G. A., WALKER, M., WILTSHIRE, S., HATTERSLEY, A. T. & MCCARTHY, M. I. 2006. Association analysis of 6,736 U.K. subjects provides replication and confirms TCF7L2 as a type 2 diabetes susceptibility gene with a substantial effect on individual risk. *Diabetes,* 55, 2640-4.

GU, Y., LEWIS, D. F. & WANG, Y. 2008. Placental Productions and Expressions of Soluble Endoglin, Soluble fms-Like Tyrosine Kinase Receptor-1, and Placental Growth Factor in Normal and Preeclamptic Pregnancies. *J Clin Endocrinol Metab,* 93, 260-266.

GULLER, S., MA, Y. Y., FU, H. H., KRIKUN, G., ABRAHAMS, V. M. & MOR, G. 2008. The placental syncytium and the pathophysiology of preeclampsia and intrauterine growth restriction: a novel assay to assess syncytial protein expression. *Ann N Y Acad Sci,* 1127, 129-33.

GUO, L., CHOUFANI, S., FERREIRA, J., SMITH, A., CHITAYAT, D., SHUMAN, C., UXA, R., KEATING, S., KINGDOM, J. & WEKSBERG, R. 2008. Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. *Dev Biol,* 320, 79-91.

HADDEN, D. R. & MCLAUGHLIN, C. 2009. Normal and abnormal maternal metabolism during pregnancy. *Semin Fetal Neonatal Med,* 14, 66-71.

HAERI, S., KHOURY, J., KOVILAM, O. & MIODOVNIK, M. 2008. The association of intrauterine growth abnormalities in women with type 1 diabetes mellitus complicated by vasculopathy. *Am J Obstet Gynecol,* 199, 278 e1-5.

HAMMING, K. S., SOLIMAN, D., MATEMISZ, L. C., NIAZI, O., LANG, Y., GLOYN, A. L. & LIGHT, P. E. 2009. Coexpression of the type 2 diabetes susceptibility gene variants KCNJ11 E23K and ABCC8 S1369A alter the ATP and sulfonylurea sensitivities of the ATP-sensitive K(+) channel. *Diabetes,* 58, 2419-24.

HAN, X., LUO, Y., REN, Q., ZHANG, X., WANG, F., SUN, X., ZHOU, X. & JI, L. 2010. Implication of genetic variants near SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, FTO, TCF2, KCNQ1, and WFS1 in Type 2 Diabetes in a Chinese population. *BMC Medical Genetics,* 11, 81.

HARRIS, L. K., KEOGH, R. J., WAREING, M., BAKER, P. N., CARTWRIGHT, J. E., APLIN, J. D. & WHITLEY, G. S. 2006. Invasive trophoblasts stimulate vascular smooth muscle cell apoptosis by a fas ligand-dependent mechanism. *Am J Pathol,* 169, 1863-74.

HARRIS, L. K., SMITH, S. D., KEOGH, R. J., JONES, R. L., BAKER, P. N., KNOFLER, M., CARTWRIGHT, J. E., WHITLEY, G. S. & APLIN, J. D. 2010. Trophoblast- and vascular smooth muscle cell-derived MMP-12 mediates elastolysis during uterine spiral artery remodeling. *Am J Pathol,* 177, 2103-15.

HARRISON, G. A., HUMPHREY, K. E., JONES, N., BADENHOP, R., GUO, G., ELAKIS, G., KAYE, J. A., TURNER, R. J., GREHAN, M., WILTON, A. N., BRENNECKE, S. P. & COOPER, D. W. 1997. A genomewide linkage study of preeclampsia/eclampsia reveals evidence for a candidate region on 4q. *Am J Hum Genet,* 60, 1158-67.

HARTLEY, J. L., TEMPLE, G. F. & BRASCH, M. A. 2000. DNA cloning using in vitro site-specific recombination. *Genome Res,* 10, 1788-95.

HATTERSLEY, A. T., BEARDS, F., BALLANTYNE, E., APPLETON, M., HARVEY, R. & ELLARD, S. 1998. Mutations in the glucokinase gene of the fetus result in reduced birth weight. *Nat Genet,* 19, 268-270.

HAY, W. W. 2006. Recent observations on the regulation of fetal metabolism by glucose. *The Journal of Physiology,* 572, 17-24.

HAYASHI, T., IWAMOTO, Y., KAKU, K., HIROSE, H. & MAEDA, S. 2007. Replication study for the association of TCF7L2 with susceptibility to type 2 diabetes in a Japanese population. *Diabetologia,* 50, 980-4.

HE, S., SILVEIRA, A., HAMSTEN, A., BLOMBACK, M. & BREMME, K. 1999. Haemostatic, endothelial and lipoprotein parameters and blood pressure levels in women with a history of preeclampsia. *Thromb Haemost,* 81, 538-42.

HELGADOTTIR, A., THORLEIFSSON, G., MAGNUSSON, K. P., GRETARSDOTTIR, S., STEINTHORSDOTTIR, V., MANOLESCU, A., JONES, G. T., RINKEL, G. J., BLANKENSTEIJN, J. D., RONKAINEN, A., JAASKELAINEN, J. E., KYO, Y., LENK, G. M., SAKALIHASAN, N., KOSTULAS, K., GOTTSATER, A., FLEX, A., STEFANSSON, H., HANSEN, T., ANDERSEN, G., WEINSHEIMER, S., BORCH-JOHNSEN, K., JORGENSEN, T., SHAH, S. H., QUYYUMI, A. A., GRANGER, C. B., REILLY, M. P., AUSTIN, H., LEVEY, A. I., VACCARINO, V., PALSDOTTIR, E., WALTERS, G. B., JONSDOTTIR, T., SNORRADOTTIR, S., MAGNUSDOTTIR, D., GUDMUNDSSON, G., FERRELL, R. E., SVEINBJORNSDOTTIR, S., HERNESNIEMI, J., NIEMELA, M., LIMET, R., ANDERSEN, K., SIGURDSSON, G., BENEDIKTSSON, R., VERHOEVEN, E. L., TEIJINK, J. A., GROBBEE, D. E., RADER, D. J., COLLIER, D. A., PEDERSEN, O., POLA, R., HILLERT, J., LINDBLAD, B., VALDIMARSSON, E. M., MAGNADOTTIR, H. B., WIJMENGA, C., TROMP, G., BAAS, A. F., RUIGROK, Y. M., VAN RIJ, A. M., KUIVANIEMI, H., POWELL, J. T., MATTHIASSON, S. E., GULCHER, J. R., THORGEIRSSON, G., KONG, A., THORSTEINSDOTTIR, U. & STEFANSSON, K. 2008. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet,* 40, 217-24.

HELGADOTTIR, A., THORLEIFSSON, G., MANOLESCU, A., GRETARSDOTTIR, S., BLONDAL, T., JONASDOTTIR, A., JONASDOTTIR, A., SIGURDSSON, A., BAKER, A., PALSSON, A., MASSON, G., GUDBJARTSSON, D. F., MAGNUSSON, K. P., ANDERSEN, K., LEVEY, A. I., BACKMAN, V. M., MATTHIASDOTTIR, S., JONSDOTTIR, T., PALSSON, S., EINARSDOTTIR, H., GUNNARSDOTTIR, S., GYLFASON, A., VACCARINO, V., HOOPER, W. C., REILLY, M. P., GRANGER, C. B., AUSTIN, H., RADER, D. J., SHAH, S. H., QUYYUMI, A. A., GULCHER, J. R., THORGEIRSSON, G., THORSTEINSDOTTIR, U., KONG, A. & STEFANSSON, K. 2007. A Common Variant on Chromosome 9p21 Affects the Risk of Myocardial Infarction. *Science,* 316, 1491-1493.

HELGASON, A., PALSSON, S., THORLEIFSSON, G., GRANT, S. F., EMILSSON, V., GUNNARSDOTTIR, S., ADEYEMO, A., CHEN, Y., CHEN, G., REYNISDOTTIR, I., BENEDIKTSSON, R., HINNEY, A.,

HANSEN, T., ANDERSEN, G., BORCH-JOHNSEN, K., JORGENSEN, T., SCHAFER, H., FARUQUE, M., DOUMATEY, A., ZHOU, J., WILENSKY, R. L., REILLY, M. P., RADER, D. J., BAGGER, Y., CHRISTIANSEN, C., SIGURDSSON, G., HEBEBRAND, J., PEDERSEN, O., THORSTEINSDOTTIR, U., GULCHER, J. R., KONG, A., ROTIMI, C. & STEFANSSON, K. 2007. Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet,* 39, 218-25.

HENDRIX, N. & BERGHELLA, V. 2008. Non-placental causes of intrauterine growth restriction. *Semin Perinatol,* 32, 161-5.

HERTEL, K. J. 2008. Combinatorial control of exon recognition. *J Biol Chem,* 283, 1211-5.

HIBINO, H., INANOBE, A., FURUTANI, K., MURAKAMI, S., FINDLAY, I. & KURACHI, Y. 2010. Inwardly Rectifying Potassium Channels: Their Structure, Function, and Physiological Roles. *Physiological Reviews,* 90, 291-366.

HOFMANN, G. E. & ABRAMOWICZ, J. S. 1990. Epidermal growth factor (EGF) concentrations in amniotic fluid and maternal urine during pregnancy. *Acta Obstet Gynecol Scand,* 69, 217-21.

HOFMANN, G. E., SCOTT, R. T., JR., BERGH, P. A. & DELIGDISCH, L. 1991. Immunohistochemical localization of epidermal growth factor in human endometrium, decidua, and placenta. *J Clin Endocrinol Metab,* 73, 882-887.

HOLSTE, D. & OHLER, U. 2008. Strategies for Identifying RNA Splicing Regulatory Motifs and Predicting Alternative Splicing Events. *PLoS Computational Biology,* 4, e21.

HONKA, M. J., VANTTINEN, M., IOZZO, P., VIRTANEN, K. A., LAUTAMAKI, R., HALLSTEN, K., BORRA, R. J., TAKALA, T., VILJANEN, A. P., KEMPPAINEN, J., PIHLAJAMAKI, J., KNUUTI, J., NUUTILA, P. & LAAKSO, M. 2009. The Pro12Ala polymorphism of the PPARgamma2 gene is associated with hepatic glucose uptake during hyperinsulinemia in subjects with type 2 diabetes mellitus. *Metabolism,* 58, 541-6.

HOOGENDOORN, B., COLEMAN, S. L., GUY, C. A., SMITH, K., BOWEN, T., BUCKLAND, P. R. & O'DONOVAN, M. C. 2003. Functional analysis of human promoter polymorphisms. *Hum Mol Genet,* 12, 2249-54.

HOUSE, A. E. & LYNCH, K. W. 2008. Regulation of alternative splicing: more than just the ABCs. *J Biol Chem,* 283, 1217-21.

HOWARTH, C., GAZIS, A. & JAMES, D. 2007. Associations of Type 1 diabetes mellitus, maternal vascular disease and complications of pregnancy. *Diabet Med,* 24, 1229-34.

HUBEL, C. A., POWERS, R. W., SNAEDAL, S., GAMMILL, H. S., NESS, R. B., ROBERTS, J. M. & ARNGRIMSSON, R. 2008. C-reactive protein is elevated 30 years after eclamptic pregnancy. *Hypertension,* 51, 1499-505.

HUMPHREY, R. G., SONNENBERG-HIRCHE, C., SMITH, S. D., HU, C., BARTON, A., SADOVSKY, Y. & NELSON, D. M. 2008. Epidermal Growth Factor Abrogates Hypoxia-Induced Apoptosis in Cultured Human Trophoblasts through Phosphorylation of BAD Serine 112. *Endocrinology,* 149, 2131-2137.

HUPPERTZ, B. 2008. Placental origins of preeclampsia: challenging the current hypothesis. *Hypertension,* 51, 970-5.
HUPPERTZ, B., GAUSTER, M., ORENDI, K., KONIG, J. & MOSER, G. 2009. Oxygen as modulator of trophoblast invasion. *J Anat,* 215, 14-20.
HYPPONEN, E., SMITH, G. D. & POWER, C. 2003. Parental diabetes and birth weight of offspring: intergenerational cohort study. *BMJ,* 326, 19-20.
ICHIBA, H., FUJIMURA, M. & TAKEUCHI, T. 1992. Levels of epidermal growth factor in human cord blood. *Biol Neonate,* 61, 302-7.
IOANNIDIS, J. P., NTZANI, E. E., TRIKALINOS, T. A. & CONTOPOULOS-IOANNIDIS, D. G. 2001. Replication validity of genetic association studies. *Nat Genet,* 29, 306-9.
ISHIHARA, N., MATSUO, H., MURAKOSHI, H., LAOAG-FERNANDEZ, J. B., SAMOTO, T. & MARUO, T. 2002. Increased apoptosis in the syncytiotrophoblast in human term placentas complicated by either preeclampsia or intrauterine growth retardation. *Am J Obstet Gynecol,* 186, 158-66.
JAMES, J. L., WHITLEY, G. S. & CARTWRIGHT, J. E. 2010. Pre-eclampsia: fitting together the placental, immune and cardiovascular pieces. *J Pathol,* 221, 363-78.
JAQUET, D., TREGOUET, D. A., GODEFROY, T., NICAUD, V., CHEVENNE, D., TIRET, L., CZERNICHOW, P. & LEVY-MARCHAL, C. 2002. Combined effects of genetic and environmental factors on insulin resistance associated with reduced fetal growth. *Diabetes,* 51, 3473-8.
JARINOVA, O., STEWART, A. F. R., ROBERTS, R., WELLS, G., LAU, P., NAING, T., BUERKI, C., MCLEAN, B. W., COOK, R. C., PARKER, J. S. & MCPHERSON, R. 2009. Functional Analysis of the Chromosome 9p21.3 Coronary Artery Disease Risk Locus. *Arterioscler Thromb Vasc Biol,* 29, 1671-1677.
JAUNIAUX, E., WATSON, A. L., HEMPSTOCK, J., BAO, Y. P., SKEPPER, J. N. & BURTON, G. J. 2000. Onset of maternal arterial blood flow and placental oxidative stress. A possible factor in human early pregnancy failure. *Am J Pathol,* 157, 2111-22.
JENSEN, C. B., STORGAARD, H., DELA, F., HOLST, J. J., MADSBAD, S. & VAAG, A. A. 2002. Early differential defects of insulin secretion and action in 19-year-old caucasian men who had low birth weight. *Diabetes,* 51, 1271-80.
JIN, T. & LIU, L. 2008. The Wnt signaling pathway effector TCF7L2 and type 2 diabetes mellitus. *Mol Endocrinol,* 22, 2383-92.
JOHNSTONE, E. D., SIBLEY, C. P., LOWEN, B. & GUILBERT, L. J. 2005. Epidermal growth factor stimulation of trophoblast differentiation requires MAPK11/14 (p38 MAP kinase) activation. *Biol Reprod,* 73, 1282-8.
JONES, H. N., POWELL, T. L. & JANSSON, T. 2007. Regulation of placental nutrient transport--a review. *Placenta,* 28, 763-74.
KAAJA, R., TIKKANEN, M. J., VIINIKKA, L. & YLIKORKALA, O. 1995. Serum lipoproteins, insulin, and urinary prostanoid metabolites in normal and hypertensive pregnant women. *Obstet Gynecol,* 85, 353-6.
KAMEI, Y., TSUTSUMI, O., YAMAKAWA, A., OKA, Y., TAKETANI, Y. & IMAKI, J. 1999. Maternal Epidermal Growth Factor Deficiency Causes Fetal Hypoglycemia and Intrauterine Growth Retardation in Mice:

Possible Involvement of Placental Glucose Transporter GLUT3 Expression. *Endocrinology,* 140, 4236-4243.

KARIMI, M., DEPICKER, A. & HILSON, P. 2007. Recombinational Cloning with Plant Gateway Vectors. *Plant Physiology,* 145, 1144-1154.

KARIMI, M., GOLDIE, L. C., CRUICKSHANK, M. N., MOSES, E. K. & ABRAHAM, L. J. 2009. A critical assessment of the factors affecting reporter gene assays for promoter SNP function: a reassessment of -308 TNF polymorphism function using a novel integrated reporter system. *Eur J Hum Genet,* 17, 1454-62.

KASHIMA, T., RAO, N., DAVID, C. J. & MANLEY, J. L. 2007. hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. *Hum Mol Genet,* 16, 3149-59.

KAUFMANN, P., BLACK, S. & HUPPERTZ, B. 2003. Endovascular trophoblast invasion: implications for the pathogenesis of intrauterine growth retardation and preeclampsia. *Biol Reprod,* 69, 1-7.

KHAN, K. S., WOJDYLA, D., SAY, L., GULMEZOGLU, A. M. & VAN LOOK, P. F. 2006. WHO analysis of causes of maternal death: a systematic review. *Lancet,* 367, 1066-74.

KILPELAINEN, T. O., DEN HOED, M., ONG, K. K., GRONTVED, A., BRAGE, S., CONSORTIUM, E. G., JAMESON, K., COOPER, C., KHAW, K. T., EKELUND, U., WAREHAM, N. J. & LOOS, R. J. 2011. Obesity-susceptibility loci have a limited influence on birth weight: a meta-analysis of up to 28,219 individuals. *Am J Clin Nutr,* 93, 851-60.

KOO, B. K., CHO, Y. M., PARK, B. L., CHEONG, H. S., SHIN, H. D., JANG, H. C., KIM, S. Y., LEE, H. K. & PARK, K. S. 2007. Polymorphisms of KCNJ11 (Kir6.2 gene) are associated with Type 2 diabetes and hypertension in the Korean population. *Diabet Med,* 24, 178-86.

KOZAK, M. 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene,* 361, 13-37.

KUPFERMINC, M. J. 2003. Thrombophilia and pregnancy. *Reprod Biol Endocrinol,* 1, 111.

LA BATIDE-ALANORE, A., TREGOUET, D. A., JAQUET, D., BOUYER, J. & TIRET, L. 2002. Familial aggregation of fetal growth restriction in a French cohort of 7,822 term births between 1971 and 1985. *Am J Epidemiol,* 156, 180-7.

LAASANEN, J., HEINONEN, S., HILTUNEN, M., MANNERMAA, A. & LAAKSO, M. 2002. Polymorphism in the peroxisome proliferator-activated receptor-gamma gene in women with preeclampsia. *Early Hum Dev,* 69, 77-82.

LADINES-LLAVE, C. A., MARUO, T., MANALO, A. S. & MOCHIZUKI, M. 1991. Cytologic localization of epidermal growth factor and its receptor in developing human placenta varies over the course of pregnancy. *Am J Obstet Gynecol,* 165, 1377-82.

LAIVUORI, H., LAHERMO, P., OLLIKAINEN, V., WIDEN, E., HAIVA-MALLINEN, L., SUNDSTROM, H., LAITINEN, T., KAAJA, R., YLIKORKALA, O. & KERE, J. 2003. Susceptibility loci for preeclampsia on chromosomes 2p25 and 9p13 in Finnish families. *Am J Hum Genet,* 72, 168-77.

LAMARCA, H. L., DASH, P. R., VISHNUTHEVAN, K., HARVEY, E., SULLIVAN, D. E., MORRIS, C. A. & WHITLEY, G. S. 2008. Epidermal growth factor-stimulated extravillous cytotrophoblast motility

is mediated by the activation of PI3-K, Akt and both p38 and p42/44 mitogen-activated protein kinases. *Hum Reprod,* 23, 1733-41.

LANDON, M. B., HAUTH, J. C., LEVENO, K. J., SPONG, C. Y., LEINDECKER, S., VARNER, M. W., MOAWAD, A. H., CARITIS, S. N., HARPER, M., WAPNER, R. J., SOROKIN, Y., MIODOVNIK, M., CARPENTER, M., PEACEMAN, A. M., O'SULLIVAN, M. J., SIBAI, B., LANGER, O., THORP, J. M., RAMIN, S. M., MERCER, B. M. & GABBE, S. G. 2004. Maternal and perinatal outcomes associated with a trial of labor after prior cesarean delivery. *N Engl J Med,* 351, 2581-9.

LANUTI, M., LIU, G., GOODWIN, J. M., ZHAI, R., FUCHS, B. C., ASOMANING, K., SU, L., NISHIOKA, N. S., TANABE, K. K. & CHRISTIANI, D. C. 2008. A functional epidermal growth factor (EGF) polymorphism, EGF serum levels, and esophageal adenocarcinoma risk and outcome. *Clin Cancer Res,* 14, 3216-22.

LAPPALAINEN, T., KOLEHMAINEN, M., SCHWAB, U. S., TOLPPANEN, A. M., STANCAKOVA, A., LINDSTROM, J., ERIKSSON, J. G., KEINANEN-KIUKAANNIEMI, S., AUNOLA, S., ILANNE-PARIKKA, P., HERDER, C., KOENIG, W., GYLLING, H., KOLB, H., TUOMILEHTO, J., KUUSISTO, J. & UUSITUPA, M. 2010. Association of the FTO gene variant (rs9939609) with cardiovascular disease in men with abnormal glucose metabolism - The Finnish Diabetes Prevention Study. *Nutr Metab Cardiovasc Dis.*

LAPPAS, M., PERMEZEL, M., GEORGIOU, H. M. & RICE, G. E. 2002. Regulation of proinflammatory cytokines in human gestational tissues by peroxisome proliferator-activated receptor-gamma: effect of 15-deoxy-Delta(12,14)-PGJ(2) and troglitazone. *J Clin Endocrinol Metab,* 87, 4667-72.

LAPPAS, M., PERMEZEL, M. & RICE, G. E. 2006. 15-Deoxy-Δ12,14-Prostaglandin J2 and Troglitazone Regulation of the Release of Phospholipid Metabolites, Inflammatory Cytokines and Proteases from Human Gestational Tissues. *Placenta,* 27, 1060-1072.

LAUENBORG, J., GRARUP, N., DAMM, P., BORCH-JOHNSEN, K., JORGENSEN, T., PEDERSEN, O. & HANSEN, T. 2009. Common Type 2 Diabetes Risk Gene Variants Associate with Gestational Diabetes. *J Clin Endocrinol Metab,* 94, 145-150.

LAUKKANEN, O., PIHLAJAMAKI, J., LINDSTROM, J., ERIKSSON, J., VALLE, T. T., HAMALAINEN, H., ILANNE-PARIKKA, P., KEINANEN-KIUKAANNIEMI, S., TUOMILEHTO, J., UUSITUPA, M. & LAAKSO, M. 2004. Polymorphisms of the SUR1 (ABCC8) and Kir6.2 (KCNJ11) genes predict the conversion from impaired glucose tolerance to type 2 diabetes. The Finnish Diabetes Prevention Study. *J Clin Endocrinol Metab,* 89, 6286-90.

LAW, A. Y. S., YEUNG, B. H. Y., CHING, L. Y. & WONG, C. K. C. 2011. Sp1 is a transcription repressor to stanniocalcin-1 expression in TSA-treated human colon cancer cells, HT29. *Journal of Cellular Biochemistry,* n/a-n/a.

LAWLOR, D. A., DAVEY SMITH, G. & EBRAHIM, S. 2002. Birth weight of offspring and insulin resistance in late adulthood: cross sectional survey. *BMJ,* 325, 359.

LEE, M. H., JEON, Y. J., LEE, S. M., PARK, M. H., JUNG, S. C. & KIM, Y. J. 2010. Placental gene expression is related to glucose metabolism and

fetal cord blood levels of insulin and insulin-like growth factors in intrauterine growth restriction. *Early Hum Dev,* 86, 45-50.

LEEMAN, L. & FONTAINE, P. 2008. Hypertensive disorders of pregnancy. *Am Fam Physician,* 78, 93-100.

LEFF, M., ORLEANS, M., HAVERKAMP, A. D., BARON, A. E., ALDERMAN, B. W. & FREEDMAN, W. L. 1992. The association of maternal low birthweight and infant low birthweight in a racially mixed population. *Paediatr Perinat Epidemiol,* 6, 51-61.

LEGRY, V., COTTEL, D., FERRIERES, J., ARVEILER, D., ANDRIEUX, N., BINGHAM, A., WAGNER, A., RUIDAVETS, J. B., DUCIMETIERE, P., AMOUYEL, P. & MEIRHAEGHE, A. 2009. Effect of an FTO polymorphism on fat mass, obesity, and type 2 diabetes mellitus in the French MONICA Study. *Metabolism,* 58, 971-5.

LEVINE, R. J., MAYNARD, S. E., QIAN, C., LIM, K. H., ENGLAND, L. J., YU, K. F., SCHISTERMAN, E. F., THADHANI, R., SACHS, B. P., EPSTEIN, F. H., SIBAI, B. M., SUKHATME, V. P. & KARUMANCHI, S. A. 2004. Circulating angiogenic factors and the risk of preeclampsia. *N Engl J Med,* 350, 672-83.

LI, L. & DAVIE, J. R. 2010. The role of Sp1 and Sp3 in normal and cancer cell biology. *Ann Anat,* 192, 275-83.

LI, R. H. & ZHUANG, L. Z. 1997. The effects of growth factors on human normal placental cytotrophoblast cell proliferation. *Hum. Reprod.,* 12, 830-834.

LIBBY, G. M., D. J. MCEWAN, N. F. GREENE, S. A. 2007. Pre-eclampsia and the later development of type 2 diabetes in mothers and their children: an intergenerational study from the Walker cohort. *Diabetologica,* 50, :523–530.

LIE, R. T., RASMUSSEN, S., BRUNBORG, H., GJESSING, H. K., LIE-NIELSEN, E. & IRGENS, L. M. 1998. Fetal and maternal contributions to risk of pre-eclampsia: population based study. *BMJ,* 316, 1343-7.

LINDGREN, C. M., MAHTANI, M. M., WIDEN, E., MCCARTHY, M. I., DALY, M. J., KIRBY, A., REEVE, M. P., KRUGLYAK, L., PARKER, A., MEYER, J., ALMGREN, P., LEHTO, M., KANNINEN, T., TUOMI, T., GROOP, L. C. & LANDER, E. S. 2002. Genomewide search for type 2 diabetes mellitus susceptibility loci in Finnish families: the Botnia study. *Am J Hum Genet,* 70, 509-16.

LINDQVIST, P., GRENNERT, L. & MARSAL, K. 1999. Epidermal growth factor in maternal urine--a predictor of intrauterine growth restriction? *Early Hum Dev,* 56, 143-50.

LINDSAY, R. S., DABELEA, D., ROUMAIN, J., HANSON, R. L., BENNETT, P. H. & KNOWLER, W. C. 2000. Type 2 diabetes and low birth weight: the role of paternal inheritance in the association of low birth weight and diabetes. *Diabetes,* 49, 445-9.

LITHELL, H. O., MCKEIGUE, P. M., BERGLUND, L., MOHSEN, R., LITHELL, U. B. & LEON, D. A. 1996. Relation of size at birth to non-insulin dependent diabetes and insulin concentrations in men aged 50-60 years. *BMJ,* 312, 406-10.

LIU, F. & GONG, C. X. 2008. Tau exon 10 alternative splicing and tauopathies. *Mol Neurodegener,* 3, 8.

LIU, H. X., ZHANG, M. & KRAINER, A. R. 1998. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes Dev,* 12, 1998-2012.

LOPEZ-BERMEJO, A., PETRY, C. J., DIAZ, M., SEBASTIANI, G., DE ZEGHER, F., DUNGER, D. B. & IBANEZ, L. 2008. The association between the FTO gene and fat mass in humans develops by the postnatal age of two weeks. *J Clin Endocrinol Metab,* 93, 1501-5.

LOUKOVAARA, M., LEINONEN, P., TERAMO, K., ANDERSSON, S., ALFTHAN, H. & STENMAN, U.-H. 2004. Diabetic Pregnancy Associated With Increased Epidermal Growth Factor in Cord Serum at Term. *Obstet Gynecol,* 103, 240-244.

LOWE, S. A., BROWN, M. A., DEKKER, G. A., GATT, S., MCLINTOCK, C. K., MCMAHON, L. P., MANGOS, G., MOORE, M. P., MULLER, P., PAECH, M. & WALTERS, B. 2009. Guidelines for the management of hypertensive disorders of pregnancy 2008. *Aust N Z J Obstet Gynaecol,* 49, 242-6.

LUNDE, A., MELVE, K. K., GJESSING, H. K., SKJAERVEN, R. & IRGENS, L. M. 2007. Genetic and Environmental Influences on Birth Weight, Birth Length, Head Circumference, and Gestational Age by Use of Population-based Parent-Offspring Data.

LUNDSTAM, U., HAGG, U., SVERRISDOTTIR, Y. B., SVENSSON, L. E. & GAN, L. M. 2007. Epidermal growth factor levels are related to diastolic blood pressure and carotid artery stiffness. *Scand Cardiovasc J,* 41, 308-12.

LUNGHI, L., FERRETTI, M. E., MEDICI, S., BIONDI, C. & VESCE, F. 2007. Control of human trophoblast function. *Reprod Biol Endocrinol,* 5, 6.

LUO, T. H., ZHAO, Y., LI, G., YUAN, W. T., ZHAO, J. J., CHEN, J. L., HUANG, W. & LUO, M. 2001. A genome-wide search for type II diabetes susceptibility genes in Chinese Hans. *Diabetologia,* 44, 501-6.

LYKKE, J. A., LANGHOFF-ROOS, J., SIBAI, B. M., FUNAI, E. F., TRICHE, E. W. & PAIDAS, M. J. 2009. Hypertensive pregnancy disorders and subsequent cardiovascular morbidity and type 2 diabetes mellitus in the mother. *Hypertension,* 53, 944-51.

LYSSENKO, V., ALMGREN, P., ANEVSKI, D., ORHO-MELANDER, M., SJOGREN, M., SALORANTA, C., TUOMI, T. & GROOP, L. 2005. Genetic prediction of future type 2 diabetes. *PLoS Med,* 2, e345.

LYSSENKO, V., JONSSON, A., ALMGREN, P., PULIZZI, N., ISOMAA, B., TUOMI, T., BERGLUND, G., ALTSHULER, D., NILSSON, P. & GROOP, L. 2008. Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med,* 359, 2220-32.

LYSSENKO, V., LUPI, R., MARCHETTI, P., DEL GUERRA, S., ORHO-MELANDER, M., ALMGREN, P., SJOGREN, M., LING, C., ERIKSSON, K. F., LETHAGEN, A. L., MANCARELLA, R., BERGLUND, G., TUOMI, T., NILSSON, P., DEL PRATO, S. & GROOP, L. 2007. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *J Clin Invest,* 117, 2155-63.

MA, L., WANG, A. H., FRIEDA, L., HE, H. Y., MA, G. X., WANG, H. W. & LIN, L. M. 2009. Effect of epidermal growth factor and dexamethasone on fetal rat lung development. *Chin Med J (Engl),* 122, 2013-6.

MAGEE, L., ME, H. & MOUTQUIN JM, E. A. 2008. Diagnosis, evaluation and management of the hypertensive disorders of pregnancy. *J Obstet Gynaecol Can,* 30, suppl: 1-48.

MAGNUS, P., BAKKETEIG, L. S. & HOFFMAN, H. 1997. Birth weight of relatives by maternal tendency to repeat small-for-gestational-age (SGA) births in successive pregnancies. *Acta Obstet Gynecol Scand Suppl,* 165, 35-8.

MAGNUS, P., GJESSING, H. K., SKRONDAL, A. & SKJAERVEN, R. 2001. Paternal contribution to birth weight. *J Epidemiol Community Health,* 55, 873-7.

MAGNUS, P., IRGENS, L. M., HAUG, K., NYSTAD, W., SKJAERVEN, R. & STOLTENBERG, C. 2006. Cohort profile: the Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol,* 35, 1146-50.

MANNHOLD, R. 2004. KATP channel openers: Structure-activity relationships and therapeutic potential. *Medicinal Research Reviews,* 24, 213-266.

MARUO, T., MATSUO, H., MURATA, K. & MOCHIZUKI, M. 1992. Gestational age-dependent dual action of epidermal growth factor on human placenta early in gestation. *J Clin Endocrinol Metab,* 75, 1362-1367.

MARUO, T., MATSUO, H., OISHI, T., HAYASHI, M., NISHINO, R. & MOCHIZUKI, M. 1987. Induction of differentiated trophoblast function by epidermal growth factor: relation of immunohistochemically detected cellular epidermal growth factor receptor levels. *J Clin Endocrinol Metab,* 64, 744-50.

MARUO, T., MATSUO, H., OTANI, T. & MOCHIZUKI, M. 1995. Role of epidermal growth factor (EGF) and its receptor in the development of the human placenta. *Reprod Fertil Dev,* 7, 1465-70.

MASUD, S. & YE, S. 2003. Effect of the peroxisome proliferator activated receptor-gamma gene Pro12Ala variant on body mass index: a meta-analysis. *J Med Genet,* 40, 773-80.

MASUGI, J., TAMORI, Y., MORI, H., KOIKE, T. & KASUGA, M. 2000. Inhibitory effect of a proline-to-alanine substitution at codon 12 of peroxisome proliferator-activated receptor-gamma 2 on thiazolidinedione-induced adipogenesis. *Biochem Biophys Res Commun,* 268, 178-82.

MASUYAMA, H., HIRAMATSU, Y. & KUDO, T. 1996. Effect of epidermal growth factor on placental amino acid transport and regulation of epidermal growth factor receptor expression of hepatocyte in rat. *J Perinat Med,* 24, 213-20.

MAYHEW, T. M., WIJESEKARA, J., BAKER, P. N. & ONG, S. S. 2004. Morphometric evidence that villous development and fetoplacental angiogenesis are compromised by intrauterine growth restriction but not by pre-eclampsia. *Placenta,* 25, 829-33.

MAYNARD, S. E., MIN, J. Y., MERCHAN, J., LIM, K. H., LI, J., MONDAL, S., LIBERMANN, T. A., MORGAN, J. P., SELLKE, F. W., STILLMAN, I. E., EPSTEIN, F. H., SUKHATME, V. P. & KARUMANCHI, S. A. 2003. Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *J Clin Invest,* 111, 649-58.

MCCLELLAND, S., SHRIVASTAVA, R. & MEDH, J. D. 2009. Regulation of Translational Efficiency by Disparate 5' UTRs of PPARgamma Splice Variants. *PPAR Res,* 2009, 193413.

MCINTIRE, D. D., BLOOM, S. L., CASEY, B. M. & LEVENO, K. J. 1999. Birth weight in relation to morbidity and mortality among newborn infants. *N Engl J Med,* 340, 1234-8.

MCPHERSON, R., PERTSEMLIDIS, A., KAVASLAR, N., STEWART, A., ROBERTS, R., COX, D. R., HINDS, D. A., PENNACCHIO, L. A., TYBJAERG-HANSEN, A., FOLSOM, A. R., BOERWINKLE, E., HOBBS, H. H. & COHEN, J. C. 2007. A Common Allele on Chromosome 9 Associated with Coronary Heart Disease. *Science,* 316, 1488-1491.

MEI, H., CHEN, W., SRINIVASAN, S. R., JIANG, F., SCHORK, N., MURRAY, S., SMITH, E., SO, J. D. & BERENSON, G. S. 2010. FTO influences on longitudinal BMI over childhood and adulthood and modulation on relationship between birth weight and longitudinal BMI. *Hum Genet,* 128, 589-96.

MEIRHAEGHE, A. & AMOUYEL, P. 2004. Impact of genetic variation of PPARgamma in humans. *Mol Genet Metab,* 83, 93-102.

MEIRHAEGHE, A., COTTEL, D., AMOUYEL, P. & DALLONGEVILLE, J. 2005. Association Between Peroxisome Proliferator-Activated Receptor {gamma} Haplotypes and the Metabolic Syndrome in French Men and Women. *Diabetes,* 54, 3043-3048.

MEIRHAEGHE, A., FAJAS, L., GOUILLEUX, F., COTTEL, D., HELBECQUE, N., AUWERX, J. & AMOUYEL, P. 2003. A functional polymorphism in a STAT5B site of the human PPAR gamma 3 gene promoter affects height and lipid metabolism in a French population. *Arterioscler Thromb Vasc Biol,* 23, 289-94.

MEIRHAEGHE, A., FAJAS, L., HELBECQUE, N., COTTEL, D., AUWERX, J., DEEB, S. S. & AMOUYEL, P. 2000. Impact of the Peroxisome Proliferator Activated Receptor gamma2 Pro12Ala polymorphism on adiposity, lipids and non-insulin-dependent diabetes mellitus. *Int J Obes Relat Metab Disord,* 24, 195-9.

MEISSNER, T., BEINBRECH, B. & MAYATEPEK, E. 1999. Congenital hyperinsulinism: Molecular basis of a heterogeneous disease. *Hum Mutat,* 13, 351-361.

MENARD, D., ARSENAULT, P. & POTHIER, P. 1988. Biologic effects of epidermal growth factor in human fetal jejunum. *Gastroenterology,* 94, 656-63.

MERCER, T. R., DINGER, M. E. & MATTICK, J. S. 2009. Long non-coding RNAs: insights into functions. *Nat Rev Genet,* 10, 155-159.

MILNE, F., REDMAN, C., WALKER, J., BAKER, P., BRADLEY, J., COOPER, C., DE SWIET, M., FLETCHER, G., JOKINEN, M., MURPHY, D., NELSON-PIERCY, C., OSGOOD, V., ROBSON, S., SHENNAN, A., TUFFNELL, A., TWADDLE, S. & WAUGH, J. 2005. The pre-eclampsia community guideline (PRECOG): how to screen for and detect onset of pre-eclampsia in the community. *BMJ,* 330, 576-80.

MISTRY, H. D., KURLAK, L. O., WILLIAMS, P. J., RAMSAY, M. M., SYMONDS, M. E. & PIPKIN, F. B. 2010. Differential expression and distribution of placental glutathione peroxidases 1, 3 and 4 in normal and preeclamptic pregnancy. *Placenta,* 31, 401-8.

MIYASHITA, T., HARIGAI, M., HANADA, M. & REED, J. C. 1994. Identification of a p53-dependent negative response element in the bcl-2 gene. *Cancer Res,* 54, 3131-5.

MOGREN, I., HOGBERG, U., WINKVIST, A. & STENLUND, H. 1999. Familial occurrence of preeclampsia. *Epidemiology,* 10, 518-22.

MOHARAM, A., ALEEM, F., FARMAKIDES, G., SCHULMAN, H., MAGHZANGY, M., RASHED, A. & KONSTANTINOU, E. 1992. Urinary and amniotic epidermal growth factor during normal and abnormal pregnancies. A comparison based upon umbilical Doppler velocimetry. *Gynecol Endocrinol,* 6, 287-92.

MOLL, S., JONES, C., CROCKER, I., BAKER, P. & HEAZELL, A. 2007. Epidermal growth factor rescues trophoblast apoptosis induced by reactive oxygen species. *Apoptosis,* 12, 1611-1622.

MONK, D. & MOORE, G. E. 2004. Intrauterine growth restriction--genetic causes and consequences. *Seminars in Fetal and Neonatal Medicine,* 9, 371-378.

MOOK-KANAMORI, D. O., DE KORT, S. W., VAN DUIJN, C. M., UITTERLINDEN, A. G., HOFMAN, A., MOLL, H. A., STEEGERS, E. A., HOKKEN-KOELEGA, A. C. & JADDOE, V. W. 2009. Type 2 diabetes gene TCF7L2 polymorphism is not associated with fetal and postnatal growth in two birth cohort studies. *BMC Med Genet,* 10, 67.

MORI, H., IKEGAMI, H., KAWAGUCHI, Y., SEINO, S., YOKOI, N., TAKEDA, J., INOUE, I., SEINO, Y., YASUDA, K., HANAFUSA, T., YAMAGATA, K., AWATA, T., KADOWAKI, T., HARA, K., YAMADA, N., GOTODA, T., IWASAKI, N., IWAMOTO, Y., SANKE, T., NANJO, K., OKA, Y., MATSUTANI, A., MAEDA, E. & KASUGA, M. 2001. The Pro12 -->Ala substitution in PPAR-gamma is associated with resistance to development of diabetes in the general population: possible involvement in impairment of insulin secretion in individuals with type 2 diabetes. *Diabetes,* 50, 891-4.

MORRISH, D. W., DAKOUR, J., LI, H., XIAO, J., MILLER, R., SHERBURNE, R., BERDAN, R. C. & GUILBERT, L. J. 1997. In vitro cultured human term cytotrophoblast: a model for normal primary epithelial cells demonstrating a spontaneous differentiation programme that requires EGF for extensive development of syncytium. *Placenta,* 18, 577-85.

MORTENSEN, J. T., THULSTRUP, A. M., LARSEN, H., MOLLER, M. & SORENSEN, H. T. 2001. Smoking, sex of the offspring, and risk of placental abruption, placenta previa, and preeclampsia: a population-based cohort study. *Acta Obstet Gynecol Scand,* 80, 894-8.

MOSES, E. K., LADE, J. A., GUO, G., WILTON, A. N., GREHAN, M., FREED, K., BORG, A., TERWILLIGER, J. D., NORTH, R., COOPER, D. W. & BRENNECKE, S. P. 2000. A genome scan in families from Australia and New Zealand confirms the presence of a maternal susceptibility locus for pre-eclampsia, on chromosome 2. *Am J Hum Genet,* 67, 1581-5.

MOTTET, D., PIROTTE, S., LAMOUR, V., HAGEDORN, M., JAVERZAT, S., BIKFALVI, A., BELLAHCENE, A., VERDIN, E. & CASTRONOVO, V. 2009. HDAC4 represses p21(WAF1/Cip1) expression in human cancer cells through a Sp1-dependent, p53-independent mechanism. *Oncogene,* 28, 243-56.

MROCZKOWSKI, B. & REICH, M. 1993. Identification of biologically active epidermal growth factor precursor in human fluids and secretions. *Endocrinology,* 132, 417-25.

MUHLHAUSER, J., CRESCIMANNO, C., KAUFMANN, P., HOFLER, H., ZACCHEO, D. & CASTELLUCCI, M. 1993. Differentiation and proliferation patterns in human trophoblast revealed by c-erbB-2 oncogene product and EGF-R. *J Histochem Cytochem,* 41, 165-73.

MURPHY, V. E., SMITH, R., GILES, W. B. & CLIFTON, V. L. 2006. Endocrine Regulation of Human Fetal Growth: The Role of the Mother, Placenta, and Fetus. *Endocr Rev,* 27, 141-169.

MUTZE, S., RUDNIK-SCHONEBORN, S., ZERRES, K. & RATH, W. 2008. Genes and the preeclampsia syndrome. *J Perinat Med,* 36, 38-58.

NAICKER, T., KHEDUN, S. M., MOODLEY, J. & PIJNENBORG, R. 2003. Quantitative analysis of trophoblast invasion in preeclampsia. *Acta Obstet Gynecol Scand,* 82, 722-9.

NAKATSUJI, Y., NISHIO, Y., TANI, N., ADACHI, K., OHMICHI, M., HISAMOTO, K., MORISHIGE, K., KURACHI, H., TASAKA, K., MURATA, Y. & MATSUURA, N. 2003. Epidermal growth factor enhances invasive activity of BeWo choriocarcinoma cells by inducing alpha2 integrin expression. *Endocr J,* 50, 703-14.

NATIONAL HIGH BLOOD PRESSURE EDUCATION PROGRAM WORKING GROUP ON HIGH BLOOD PRESSURE IN PREGNANCY 2000. Report of the National High Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy. *American Journal of Obstetrics and Gynecology,* 183, S1-S22.

NEJATIZADEH, A., STOBDAN, T., MALHOTRA, N. & PASHA, M. A. 2008. The genetic aspects of pre-eclampsia: achievements and limitations. *Biochem Genet,* 46, 451-79.

NETCHINE, I., ROSSIGNOL, S., DUFOURG, M. N., AZZI, S., ROUSSEAU, A., PERIN, L., HOUANG, M., STEUNOU, V., ESTEVA, B., THIBAUD, N., DEMAY, M. C., DANTON, F., PETRICZKO, E., BERTRAND, A. M., HEINRICHS, C., CAREL, J. C., LOEUILLE, G. A., PINTO, G., JACQUEMONT, M. L., GICQUEL, C., CABROL, S. & LE BOUC, Y. 2007. 11p15 imprinting center region 1 loss of methylation is a common and specific cause of typical Russell-Silver syndrome: clinical scoring system and epigenetic-phenotypic correlations. *J Clin Endocrinol Metab,* 92, 3148-54.

NEWTON-CHEH, C. & HIRSCHHORN, J. N. 2005. Genetic association studies of complex traits: design and analysis issues. *Mutat Res,* 573, 54-69.

NG, M. C., TAM, C. H., LAM, V. K., SO, W. Y., MA, R. C. & CHAN, J. C. 2007. Replication and identification of novel variants at TCF7L2 associated with type 2 diabetes in Hong Kong Chinese. *J Clin Endocrinol Metab,* 92, 3733-7.

NIELSEN, E. M., HANSEN, L., CARSTENSEN, B., ECHWALD, S. M., DRIVSHOLM, T., GLUMER, C., THORSTEINSSON, B., BORCH-JOHNSEN, K., HANSEN, T. & PEDERSEN, O. 2003. The E23K variant of Kir6.2 associates with impaired post-OGTT serum insulin response and increased risk of type 2 diabetes. *Diabetes,* 52, 573-7.

NILSSON, E., SALONEN ROS, H., CNATTINGIUS, S. & LICHTENSTEIN, P. 2004. The importance of genetic and environmental effects for pre-

eclampsia and gestational hypertension: a family study. *BJOG,* 111, 200-6.

NORTH, R. A., FERRIER, C., LONG, D., TOWNEND, K. & KINCAID-SMITH, P. 1994. Uterine artery Doppler flow velocity waveforms in the second trimester for the prediction of preeclampsia and fetal growth retardation. *Obstet Gynecol,* 83, 378-86.

ODAKA, K., HIRAMATSU, Y., EGUCHI, K. & KUDO, T. 1993. Effects of epidermal growth factor on mouse fetal growth. *Asia Oceania J Obstet Gynaecol,* 19, 213-6.

OLSON, T. & TERZIC, A. 2010. Human K(ATP) channelopathies: diseases of metabolic homeostasis. *Pflügers Archiv European Journal of Physiology,* 460, 295-306.

ONG, K. K., DIDERHOLM, B., SALZANO, G., WINGATE, D., HUGHES, I. A., MACDOUGALL, J., ACERINI, C. L. & DUNGER, D. B. 2008. Pregnancy insulin, glucose, and BMI contribute to birth outcomes in nondiabetic mothers. *Diabetes Care,* 31, 2193-7.

OSLER, M., LUND, R., KRIEGBAUM, M. & ANDERSEN, A.-M. 2009. The influence of birth weight and body mass in early adulthood on early coronary heart disease risk among Danish men born in 1953. *European Journal of Epidemiology,* 24, 57-61.

OSMARK, P., HANSSON, O., JONSSON, A., RONN, T., GROOP, L. & RENSTROM, E. 2009. Unique splicing pattern of the TCF7L2 gene in human pancreatic islets. *Diabetologia,* 52, 850-4.

OSMOND, C. & BARKER, D. J. 2000. Fetal, infant, and childhood growth are predictors of coronary heart disease, diabetes, and hypertension in adult men and women. *Environ Health Perspect,* 108 Suppl 3, 545-53.

PADMANABHAN, S., WALLACE, C., MUNROE, P. B., DOBSON, R., BROWN, M., SAMANI, N., CLAYTON, D., FARRALL, M., WEBSTER, J., LATHROP, M., CAULFIELD, M., DOMINICZAK, A. F. & CONNELL, J. M. 2006. Chromosome 2p shows significant linkage to antihypertensive response in the British Genetics of Hypertension Study. *Hypertension,* 47, 603-8.

PAGANI, F., STUANI, C., TZETIS, M., KANAVAKIS, E., EFTHYMIADOU, A., DOUDOUNAKIS, S., CASALS, T. & BARALLE, F. E. 2003. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Hum Mol Genet,* 12, 1111-20.

PALMER, L. J. & CARDON, L. R. 2005. Shaking the tree: mapping complex disease genes with linkage disequilibrium. *Lancet,* 366, 1223-34.

PARK, J. H., STOFFERS, D. A., NICHOLLS, R. D. & SIMMONS, R. A. 2008. Development of type 2 diabetes following intrauterine growth retardation in rats is associated with progressive epigenetic silencing of Pdx1. *J Clin Invest,* 118, 2316-24.

PAUSOVA, Z., SYME, C., ABRAHAMOWICZ, M., XIAO, Y., LEONARD, G. T., PERRON, M., RICHER, L., VEILLETTE, S., SMITH, G. D., SEDA, O., TREMBLAY, J., HAMET, P., GAUDET, D. & PAUS, T. 2009. A common variant of the FTO gene is associated with not only increased adiposity but also elevated blood pressure in French Canadians. *Circ Cardiovasc Genet,* 2, 260-9.

PECIOSKA, S., ZILLIKENS, M. C., HENNEMAN, P., SNIJDERS, P. J., OOSTRA, B. A., VAN DUIJN, C. M. & AULCHENKO, Y. S. 2010.

Association between Type 2 Diabetes Loci and Measures of Fatness. *PLoS ONE,* 5, e8541.

PELLANDA, L. C., DUNCAN, B. B., VIGO, A., ROSE, K., FOLSOM, A. R. & ERLINGER, T. P. 2009. Low birth weight and markers of inflammation and endothelial activation in adulthood: the ARIC study. *Int J Cardiol,* 134, 371-7.

PFAB, T., PORALLA, C., RICHTER, C. M., GODES, M., SLOWINSKI, T., PRIEM, F., HALLE, H. & HOCHER, B. 2006. Fetal and maternal peroxisome proliferator-activated receptor gamma2 Pro12Ala does not influence birth weight. *Obesity (Silver Spring),* 14, 1880-5.

PICKLESIMER, A. H., JARED, H. L., MOSS, K., OFFENBACHER, S., BECK, J. D. & BOGGESS, K. A. 2008. Racial differences in C-reactive protein levels during normal pregnancy. *Am J Obstet Gynecol,* 199, 523 e1-6.

PLATZ, E. & NEWMAN, R. 2008. Diagnosis of IUGR: traditional biometry. *Semin Perinatol,* 32, 140-7.

PLUNKETT, J., BORECKI, I., MORGAN, T., STAMILIO, D. & MUGLIA, L. J. 2008. Population-based estimate of sibling risk for preterm birth, preterm premature rupture of membranes, placental abruption and pre-eclampsia. *BMC Genet,* 9, 44.

POLLHEIMER, J., LOREGGER, T., SONDEREGGER, S., SALEH, L., BAUER, S., BILBAN, M., CZERWENKA, K., HUSSLEIN, P. & KNOFLER, M. 2006. Activation of the canonical wingless/T-cell factor signaling pathway promotes invasive differentiation of human trophoblast. *Am J Pathol,* 168, 1134-47.

PONTING, C. P., OLIVER, P. L. & REIK, W. 2009. Evolution and Functions of Long Noncoding RNAs. 136, 629-641.

POSTON, L., BRILEY, A. L., SEED, P. T., KELLY, F. J. & SHENNAN, A. H. 2006. Vitamin C and vitamin E in pregnant women at risk for pre-eclampsia (VIP trial): randomised placebo-controlled trial. *Lancet,* 367, 1145-54.

POUTA, A. H., AL SOVIO U ET AL 2004 Manifestations of metabolic syndrome after hypertensive pregnancy. . *Hypertension*
43, 825–831.

PROKUNINA-OLSSON, L., WELCH, C., HANSSON, O., ADHIKARI, N., SCOTT, L. J., USHER, N., TONG, M., SPRAU, A., SWIFT, A., BONNYCASTLE, L. L., ERDOS, M. R., HE, Z., SAXENA, R., HARMON, B., KOTOVA, O., HOFFMAN, E. P., ALTSHULER, D., GROOP, L., BOEHNKE, M., COLLINS, F. S. & HALL, J. L. 2009. Tissue-specific alternative splicing of TCF7L2. *Hum Mol Genet,* 18, 3795-804.

PULIZZI, N., LYSSENKO, V., JONSSON, A., OSMOND, C., LAAKSO, M., KAJANTIE, E., BARKER, D. J., GROOP, L. C. & ERIKSSON, J. G. 2009. Interaction between prenatal growth and high-risk genotypes in the development of type 2 diabetes. *Diabetologia,* 52, 825-9.

QIU, C., WILLIAMS, M. A., LEISENRING, W. M., SORENSEN, T. K., FREDERICK, I. O., DEMPSEY, J. C. & LUTHY, D. A. 2003. Family History of Hypertension and Type 2 Diabetes in Relation to Preeclampsia Risk. *Hypertension,* 41, 408-413.

QIU, Q., YANG, M., TSANG, B. K. & GRUSLIN, A. 2004. EGF-induced trophoblast secretion of MMP-9 and TIMP-1 involves activation of both PI3K and MAPK signalling pathways. *Reproduction,* 128, 355-363.

RAABERG, L., NEXO, E., JORGENSEN, P. E., POULSEN, S. S. & JAKAB, M. 1995. Fetal effects of epidermal growth factor deficiency induced in rats by autoantibodies against epidermal growth factor. *Pediatr Res,* 37, 175-81.

RAMSAY, J. E., STEWART, F., GREER, I. A. & SATTAR, N. 2003. Microvascular dysfunction: a link between pre-eclampsia and maternal coronary heart disease. *BJOG,* 110, 1029-31.

RANDERSON-MOOR, J. A., GAUT, R., TURNER, F., WHITAKER, L., BARRETT, J. H., SILVA IDOS, S., SWERDLOW, A., BISHOP, D. T. & BISHOP, J. A. 2004. The relationship between the epidermal growth factor (EGF) 5'UTR variant A61G and melanoma/nevus susceptibility. *J Invest Dermatol,* 123, 755-9.

RAO, C. V., LI, X., TOTH, P. & LEI, Z. M. 1995. Expression of epidermal growth factor, transforming growth factor- alpha, and their common receptor genes in human umbilical cords. *J Clin Endocrinol Metab,* 80, 1012-1020.

RAY, J. G., VERMEULEN, M. J., SCHULL, M. J. & REDELMEIER, D. A. 2005. Cardiovascular health after maternal placental syndromes (CHAMPS): population-based retrospective cohort study. *Lancet,* 366, 1797-803.

REDMAN, C. W. & SARGENT, I. L. 2005. Latest advances in understanding preeclampsia. *Science,* 308, 1592-4.

REDMAN, C. W. & SARGENT, I. L. 2008. Circulating microparticles in normal pregnancy and pre-eclampsia. *Placenta,* 29 Suppl A, S73-7.

REEVES, S. & BERNSTEIN, I. 2008a. Effects of maternal tobacco-smoke exposure on fetal growth and neonatal size. *Expert Rev Obstet Gynecol,* 3, 719-730.

REEVES, S. & BERNSTEIN, I. M. 2008b. Optimal growth modeling. *Semin Perinatol,* 32, 148-53.

REIK, W., CONSTANCIA, M., FOWDEN, A., ANDERSON, N., DEAN, W., FERGUSON-SMITH, A., TYCKO, B. & SIBLEY, C. 2003. Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. *J Physiol,* 547, 35-44.

RICH-EDWARDS, J. W., STAMPFER, M. J., MANSON, J. E., ROSNER, B., HANKINSON, S. E., COLDITZ, G. A., WILLETT, W. C. & HENNEKENS, C. H. 1997. Birth weight and risk of cardiovascular disease in a cohort of women followed up since 1976. *BMJ,* 315, 396-400.

RIDKER, P. M., COOK, N. R., CHENG, S., ERLICH, H. A., LINDPAINTNER, K., PLUTZKY, J. & ZEE, R. Y. 2003. Alanine for proline substitution in the peroxisome proliferator-activated receptor gamma-2 (PPARG2) gene and the risk of incident myocardial infarction. *Arterioscler Thromb Vasc Biol,* 23, 859-63.

ROBERTS, D. J. & POST, M. D. 2008. The placenta in pre-eclampsia and intrauterine growth restriction. *J Clin Pathol,* 61, 1254-60.

ROBERTS, J. M. & HUBEL, C. A. 2009. The two stage model of preeclampsia: variations on the theme. *Placenta,* 30 Suppl A, S32-7.

RODIE, V. A., YOUNG, A., JORDAN, F., SATTAR, N., GREER, I. A. & FREEMAN, D. J. 2005. Human placental peroxisome proliferator-activated receptor delta and gamma expression in healthy pregnancy and

in preeclampsia and intrauterine growth restriction. *J Soc Gynecol Investig,* 12, 320-9.

ROMERO, R., NIEN, J. K., ESPINOZA, J., TODEM, D., FU, W., CHUNG, H., KUSANOVIC, J. P., GOTSCH, F., EREZ, O., MAZAKI-TOVI, S., GOMEZ, R., EDWIN, S., CHAIWORAPONGSA, T., LEVINE, R. J. & KARUMANCHI, S. A. 2008. A longitudinal study of angiogenic (placental growth factor) and anti-angiogenic (soluble endoglin and soluble vascular endothelial growth factor receptor-1) factors in normal pregnancy and patients destined to develop preeclampsia and deliver a small for gestational age neonate. *J Matern Fetal Neonatal Med,* 21, 9-23.

ROSENBERG, T. J., GARBERS, S., LIPKIND, H. & CHIASSON, M. A. 2005. Maternal obesity and diabetes as risk factors for adverse pregnancy outcomes: differences among 4 racial/ethnic groups. *Am J Public Health,* 95, 1545-51.

ROSS, M. G. & BEALL, M. H. 2008. Adult sequelae of intrauterine growth restriction. *Semin Perinatol,* 32, 213-8.

ROTEN, L. T., JOHNSON, M. P., FORSMO, S., FITZPATRICK, E., DYER, T. D., BRENNECKE, S. P., BLANGERO, J., MOSES, E. K. & AUSTGULEN, R. 2009. Association between the candidate susceptibility gene ACVR2A on chromosome 2q22 and pre-eclampsia in a large Norwegian population-based study (the HUNT study). *Eur J Hum Genet,* 17, 250-7.

RUI, X., TSAO, J., SCHEYS, J. O., HAMMER, G. D. & SCHIMMER, B. P. 2008. Contributions of specificity protein-1 and steroidogenic factor 1 to Adcy4 expression in Y1 mouse adrenal cells. *Endocrinology,* 149, 3668-78.

RUMBOLD, A. R., CROWTHER, C. A., HASLAM, R. R., DEKKER, G. A. & ROBINSON, J. S. 2006. Vitamins C and E and the risks of preeclampsia and perinatal complications. *N Engl J Med,* 354, 1796-806.

RUST, S., FUNKE, H. & ASSMANN, G. 1993. Mutagenically separated PCR (MS-PCR): a highly specific one step procedure for easy mutation detection. *Nucleic Acids Res,* 21, 3623-9.

SACKS, G. P., STUDENA, K., SARGENT, K. & REDMAN, C. W. 1998. Normal pregnancy and preeclampsia both produce inflammatory changes in peripheral blood leukocytes akin to those of sepsis. *Am J Obstet Gynecol,* 179, 80-6.

SALONEN ROS, H., LICHTENSTEIN, P., LIPWORTH, L. & CNATTINGIUS, S. 2000. Genetic effects on the liability of developing pre-eclampsia and gestational hypertension. *Am J Med Genet,* 91, 256-60.

SAMANI, N. J., ERDMANN, J., HALL, A. S., HENGSTENBERG, C., MANGINO, M., MAYER, B., DIXON, R. J., MEITINGER, T., BRAUND, P., WICHMANN, H. E., BARRETT, J. H., KONIG, I. R., STEVENS, S. E., SZYMCZAK, S., TREGOUET, D.-A., ILES, M. M., PAHLKE, F., POLLARD, H., LIEB, W., CAMBIEN, F., FISCHER, M., OUWEHAND, W., BLANKENBERG, S., BALMFORTH, A. J., BAESSLER, A., BALL, S. G., STROM, T. M., BRAENNE, I., GIEGER, C., DELOUKAS, P., TOBIN, M. D., ZIEGLER, A., THOMPSON, J. R., SCHUNKERT, H., THE, W. & THE CARDIOGENICS, C. 2007. Genomewide Association Analysis of Coronary Artery Disease. *N Engl J Med,* 357, 443-453.

SANDRIM, V. C., PALEI, A. C., METZGER, I. F., GOMES, V. A., CAVALLI, R. C. & TANUS-SANTOS, J. E. 2008. Nitric oxide formation is inversely related to serum levels of antiangiogenic factors soluble fms-like tyrosine kinase-1 and soluble endogline in preeclampsia. *Hypertension,* 52, 402-7.

SATTAR, N., BENDOMIR, A., BERRY, C., SHEPHERD, J., GREER, I. A. & PACKARD, C. J. 1997. Lipoprotein subfraction concentrations in preeclampsia: pathogenic parallels to atherosclerosis. *Obstet Gynecol,* 89, 403-8.

SATTAR, N. & GREER, I. A. 2002. Pregnancy complications and maternal cardiovascular risk: opportunities for intervention and screening? *BMJ,* 325, 157-60.

SATTAR, N., RAMSAY, J., CRAWFORD, L., CHEYNE, H. & GREER, I. A. 2003. Classic and novel risk factor parameters in women with a history of preeclampsia. *Hypertension,* 42, 39-42.

SAXENA, R., GIANNINY, L., BURTT, N. P., LYSSENKO, V., GIUDUCCI, C., SJOGREN, M., FLOREZ, J. C., ALMGREN, P., ISOMAA, B., ORHO-MELANDER, M., LINDBLAD, U., DALY, M. J., TUOMI, T., HIRSCHHORN, J. N., ARDLIE, K. G., GROOP, L. C. & ALTSHULER, D. 2006. Common single nucleotide polymorphisms in TCF7L2 are reproducibly associated with type 2 diabetes and reduce the insulin response to glucose in nondiabetic individuals. *Diabetes,* 55, 2890-5.

SAXENA, R., VOIGHT, B. F., LYSSENKO, V., BURTT, N. P., DE BAKKER, P. I., CHEN, H., ROIX, J. J., KATHIRESAN, S., HIRSCHHORN, J. N., DALY, M. J., HUGHES, T. E., GROOP, L., ALTSHULER, D., ALMGREN, P., FLOREZ, J. C., MEYER, J., ARDLIE, K., BENGTSSON BOSTROM, K., ISOMAA, B., LETTRE, G., LINDBLAD, U., LYON, H. N., MELANDER, O., NEWTON-CHEH, C., NILSSON, P., ORHO-MELANDER, M., RASTAM, L., SPELIOTES, E. K., TASKINEN, M. R., TUOMI, T., GUIDUCCI, C., BERGLUND, A., CARLSON, J., GIANNINY, L., HACKETT, R., HALL, L., HOLMKVIST, J., LAURILA, E., SJOGREN, M., STERNER, M., SURTI, A., SVENSSON, M., TEWHEY, R., BLUMENSTIEL, B., PARKIN, M., DEFELICE, M., BARRY, R., BRODEUR, W., CAMARATA, J., CHIA, N., FAVA, M., GIBBONS, J., HANDSAKER, B., HEALY, C., NGUYEN, K., GATES, C., SOUGNEZ, C., GAGE, D., NIZZARI, M., GABRIEL, S. B., CHIRN, G. W., MA, Q., PARIKH, H., RICHARDSON, D., RICKE, D. & PURCELL, S. 2007. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science,* 316, 1331-6.

SCHAID, D. J. 1998. Transmission disequilibrium, family controls, and great expectations. *Am J Hum Genet,* 63, 935-41.

SCHAID, D. J. 2004. Evaluating associations of haplotypes with traits. *Genet Epidemiol,* 27, 348-64.

SCHAIFF, W. T., CARLSON, M. G., SMITH, S. D., LEVY, R., NELSON, D. M. & SADOVSKY, Y. 2000. Peroxisome proliferator-activated receptor-gamma modulates differentiation of human trophoblast in a ligand-specific manner. *J Clin Endocrinol Metab,* 85, 3874-81.

SCHUNKERT, H., GOTZ, A., BRAUND, P., MCGINNIS, R., TREGOUET, D.-A., MANGINO, M., LINSEL-NITSCHKE, P., CAMBIEN, F., HENGSTENBERG, C., STARK, K., BLANKENBERG, S., TIRET, L.,

DUCIMETIERE, P., KENIRY, A., GHORI, M. J. R., SCHREIBER, S., EL MOKHTARI, N. E., HALL, A. S., DIXON, R. J., GOODALL, A. H., LIPTAU, H., POLLARD, H., SCHWARZ, D. F., HOTHORN, L. A., WICHMANN, H. E., KONIG, I. R., FISCHER, M., MEISINGER, C., OUWEHAND, W., DELOUKAS, P., THOMPSON, J. R., ERDMANN, J., ZIEGLER, A., SAMANI, N. J. & FOR THE CARDIOGENICS, C. 2008. Repeated Replication and a Prospective Meta-Analysis of the Association Between Chromosome 9p21.3 and Coronary Artery Disease. *Circulation,* 117, 1675-1684.

SCHWANSTECHER, C., MEYER, U. & SCHWANSTECHER, M. 2002. K(IR)6.2 polymorphism predisposes to type 2 diabetes by inducing overactivity of pancreatic beta-cell ATP-sensitive K(+) channels. *Diabetes,* 51, 875-9.

SCOTT, L. J., BONNYCASTLE, L. L., WILLER, C. J., SPRAU, A. G., JACKSON, A. U., NARISU, N., DUREN, W. L., CHINES, P. S., STRINGHAM, H. M., ERDOS, M. R., VALLE, T. T., TUOMILEHTO, J., BERGMAN, R. N., MOHLKE, K. L., COLLINS, F. S. & BOEHNKE, M. 2006. Association of transcription factor 7-like 2 (TCF7L2) variants with type 2 diabetes in a Finnish sample. *Diabetes,* 55, 2649-53.

SCOTT, L. J., MOHLKE, K. L., BONNYCASTLE, L. L., WILLER, C. J., LI, Y., DUREN, W. L., ERDOS, M. R., STRINGHAM, H. M., CHINES, P. S., JACKSON, A. U., PROKUNINA-OLSSON, L., DING, C. J., SWIFT, A. J., NARISU, N., HU, T., PRUIM, R., XIAO, R., LI, X. Y., CONNEELY, K. N., RIEBOW, N. L., SPRAU, A. G., TONG, M., WHITE, P. P., HETRICK, K. N., BARNHART, M. W., BARK, C. W., GOLDSTEIN, J. L., WATKINS, L., XIANG, F., SARAMIES, J., BUCHANAN, T. A., WATANABE, R. M., VALLE, T. T., KINNUNEN, L., ABECASIS, G. R., PUGH, E. W., DOHENY, K. F., BERGMAN, R. N., TUOMILEHTO, J., COLLINS, F. S. & BOEHNKE, M. 2007. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science,* 316, 1341-5.

SEMINA, E. V., DATSON, N. A., LEYSENS, N. J., ZABEL, B. U., CAREY, J. C., BELL, G. I., BITOUN, P., LINDGREN, C., STEVENSON, T., FRANTS, R. R., VAN OMMEN, G. & MURRAY, J. C. 1996. Exclusion of epidermal growth factor and high-resolution physical mapping across the Rieger syndrome locus. *Am J Hum Genet,* 59, 1288-96.

SERDAR, Z., GUR, E., COLAKOETHULLARY, M., DEVELIOETHLU, O. & SARANDOL, E. 2003. Lipid and protein oxidation and antioxidant function in women with mild and severe preeclampsia. *Arch Gynecol Obstet,* 268, 19-25.

SESTI, G., LARATTA, E., CARDELLINI, M., ANDREOZZI, F., DEL GUERRA, S., IRACE, C., GNASSO, A., GRUPILLO, M., LAURO, R., HRIBAL, M. L., PERTICONE, F. & MARCHETTI, P. 2006. The E23K variant of KCNJ11 encoding the pancreatic beta-cell adenosine 5'-triphosphate-sensitive potassium channel subunit Kir6.2 is associated with an increased risk of secondary failure to sulfonylurea in patients with type 2 diabetes. *J Clin Endocrinol Metab,* 91, 2334-9.

SHAARAWY, M. & DIDY, H. E. 1996. Thrombomodulin, plasminogen activator inhibitor type 1 (PAI-1) and fibronectin as biomarkers of endothelial damage in preeclampsia and eclampsia. *Int J Gynaecol Obstet,* 55, 135-9.

SHAAT, N., EKELUND, M., LERNMARK, A., IVARSSON, S., ALMGREN, P., BERNTORP, K. & GROOP, L. 2005. Association of the E23K polymorphism in the KCNJ11 gene with gestational diabetes mellitus. *Diabetologia,* 48, 2544-51.

SHAAT, N., LERNMARK, A., KARLSSON, E., IVARSSON, S., PARIKH, H., BERNTORP, K. & GROOP, L. 2007. A variant in the transcription factor 7-like 2 (TCF7L2) gene is associated with an increased risk of gestational diabetes mellitus. *Diabetologia,* 50, 972-9.

SHAHBAZI, M., PRAVICA, V., NASREEN, N., FAKHOURY, H., FRYER, A. A., STRANGE, R. C., HUTCHINSON, P. E., OSBORNE, J. E., LEAR, J. T., SMITH, A. G. & HUTCHINSON, I. V. 2002. Association between functional polymorphism in EGF gene and malignant melanoma. *Lancet,* 359, 397-401.

SHARMA, P., KUMAR, J., GARG, G., KUMAR, A., PATOWARY, A., KARTHIKEYAN, G., RAMAKRISHNAN, L., BRAHMACHARI, V. & SENGUPTA, S. 2008. Detection of altered global DNA methylation in coronary artery disease patients. *DNA Cell Biol,* 27, 357-65.

SHEN, H. & GREEN, M. R. 2006. RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. *Genes Dev,* 20, 1755-65.

SHIBATA, E., RAJAKUMAR, A., POWERS, R. W., LARKIN, R. W., GILMOUR, C., BODNAR, L. M., CROMBLEHOLME, W. R., NESS, R. B., ROBERTS, J. M. & HUBEL, C. A. 2005. Soluble fms-like tyrosine kinase 1 is increased in preeclampsia but not in normotensive pregnancies with small-for-gestational-age neonates: relationship to circulating placental growth factor. *J Clin Endocrinol Metab,* 90, 4895-903.

SHU, L., MATVEYENKO, A. V., KERR-CONTE, J., CHO, J. H., MCINTOSH, C. H. & MAEDLER, K. 2009. Decreased TCF7L2 protein levels in type 2 diabetes mellitus correlate with downregulation of GIP- and GLP-1 receptors and impaired beta-cell function. *Hum Mol Genet,* 18, 2388-99.

SHU, L., SAUTER, N. S., SCHULTHESS, F. T., MATVEYENKO, A. V., OBERHOLZER, J. & MAEDLER, K. 2008. Transcription factor 7-like 2 regulates beta-cell survival and function in human pancreatic islets. *Diabetes,* 57, 645-53.

SHYNG, S. & NICHOLS, C. G. 1997. Octameric stoichiometry of the KATP channel complex. *J Gen Physiol,* 110, 655-64.

SIBAI, B., DEKKER, G. & KUPFERMINC, M. 2005. Pre-eclampsia. *Lancet,* 365, 785-99.

SIBAI, B. M. 2003. Diagnosis and management of gestational hypertension and preeclampsia. *Obstet Gynecol,* 102, 181-92.

SIBLEY, C. P., COAN, P. M., FERGUSON-SMITH, A. C., DEAN, W., HUGHES, J., SMITH, P., REIK, W., BURTON, G. J., FOWDEN, A. L. & CONSTANCIA, M. 2004. Placental-specific insulin-like growth factor 2 (Igf2) regulates the diffusional exchange characteristics of the mouse placenta. *Proc Natl Acad Sci U S A,* 101, 8204-8.

SIBLEY, D. R., BENOVIC, J. L., CARON, M. G. & LEFKOWITZ, R. J. 1987. Regulation of transmembrane signaling by receptor phosphorylation. *Cell,* 48, 913-22.

SIERRA-LAGUADO, J., GARCIA, R. G., CELEDON, J., ARENAS-MANTILLA, M., PRADILLA, L. P., CAMACHO, P. A. & LOPEZ-

JARAMILLO, P. 2007. Determination of insulin resistance using the homeostatic model assessment (HOMA) and its relation with the risk of developing pregnancy-induced hypertension. *Am J Hypertens,* 20, 437-42.

SINGH, I. S., HE, J. R., CALDERWOOD, S. & HASDAY, J. D. 2002. A high affinity HSF-1 binding site in the 5'-untranslated region of the murine tumor necrosis factor-alpha gene is a transcriptional repressor. *J Biol Chem,* 277, 4981-8.

SKJAERVEN, R., VATTEN, L. J., WILCOX, A. J., RONNING, T., IRGENS, L. M. & LIE, R. T. 2005. Recurrence of pre-eclampsia across generations: exploring fetal and maternal genetic components in a population based cohort. *BMJ,* 331, 877.

SKJAERVEN, R., WILCOX, A. J. & LIE, R. T. 2002. The interval between pregnancies and the risk of preeclampsia. *N Engl J Med,* 346, 33-8.

SLADEK, R., ROCHELEAU, G., RUNG, J., DINA, C., SHEN, L., SERRE, D., BOUTIN, P., VINCENT, D., BELISLE, A., HADJADJ, S., BALKAU, B., HEUDE, B., CHARPENTIER, G., HUDSON, T. J., MONTPETIT, A., PSHEZHETSKY, A. V., PRENTKI, M., POSNER, B. I., BALDING, D. J., MEYRE, D., POLYCHRONAKOS, C. & FROGUEL, P. 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature,* 445, 881-5.

SMITH, S., FRANCIS, R., GUILBERT, L. & BAKER, P. N. 2002. Growth Factor Rescue of Cytokine Mediated Trophoblast Apoptosis. *Placenta,* 23, 322-330.

SOLOMON, C. G., CARROLL, J. S., OKAMURA, K., GRAVES, S. W. & SEELY, E. W. 1999. Higher cholesterol and insulin levels in pregnancy are associated with increased risk for pregnancy-induced hypertension. *Am J Hypertens,* 12, 276-82.

SOLOMON, C. G., GRAVES, S. W., GREENE, M. F. & SEELY, E. W. 1994. Glucose intolerance as a predictor of hypertension in pregnancy. *Hypertension,* 23, 717-21.

SORENSEN, S., VON TABOUILLOT, D., SCHIOLER, V., GREISEN, G., PETERSEN, S. & LARSEN, T. 2000. Serial measurements of serum human placental lactogen (hPL) and serial ultrasound examinations in the evaluation of fetal growth. *Early Hum Dev,* 60, 25-34.

SOUSA, A. G. P., MARQUEZINE, G. F., LEMOS, P. A., MARTINEZ, E., LOPES, N., HUEB, W. A., KRIEGER, J. E. & PEREIRA, A. C. 2009. TCF7L2 Polymorphism rs7903146 Is Associated with Coronary Artery Disease Severity and Mortality. *PLoS ONE,* 4, e7697.

SOWA, Y., ORITA, T., MINAMIKAWA-HIRANABE, S., MIZUNO, T., NOMURA, H. & SAKAI, T. 1999. Sp3, but not Sp1, Mediates the Transcriptional Activation of the p21/WAF1/Cip1 Gene Promoter by Histone Deacetylase Inhibitor. *Cancer Research,* 59, 4266-4270.

SOWERS, J. R., STANDLEY, P. R., JACOBER, S., NIYOGI, T. & SIMPSON, L. 1993. Postpartum abnormalities of carbohydrate and cellular calcium metabolism in pregnancy induced hypertension. *Am J Hypertens,* 6, 302-7.

SPEAKMAN, J. R., RANCE, K. A. & JOHNSTONE, A. M. 2008. Polymorphisms of the FTO gene are associated with variation in energy intake, but not energy expenditure. *Obesity (Silver Spring),* 16, 1961-5.

SPINDLER, K. L., NIELSEN, J. N., ORNSKOV, D., BRANDSLUND, I. & JAKOBSEN, A. 2007. Epidermal growth factor (EGF) A61G polymorphism and EGF gene expression in normal colon tissue from patients with colorectal cancer. *Acta Oncol,* 46, 1113-7.

STAUN-RAM, E., GOLDMAN, S., GABARIN, D. & SHALEV, E. 2004. Expression and importance of matrix metalloproteinase 2 and 9 (MMP-2 and -9) in human trophoblast invasion. *Reprod Biol Endocrinol,* 2, 59.

STEEGERS, E. A., VON DADELSZEN, P., DUVEKOT, J. J. & PIJNENBORG, R. 2010. Pre-eclampsia. *Lancet,* 376, 631-44.

STRATIGOPOULOS, G., PADILLA, S. L., LEDUC, C. A., WATSON, E., HATTERSLEY, A. T., MCCARTHY, M. I., ZELTSER, L. M., CHUNG, W. K. & LEIBEL, R. L. 2008. Regulation of Fto/Ftm gene expression in mice and humans. *Am J Physiol Regul Integr Comp Physiol,* 294, R1185-96.

STUMVOLL, M., GOLDSTEIN, B. J. & VAN HAEFTEN, T. W. 2005. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet,* 365, 1333-46.

STUMVOLL, M., WAHL, H. G., LOBLEIN, K., BECKER, R., MACHICAO, F., JACOB, S. & HARING, H. 2001. Pro12Ala polymorphism in the peroxisome proliferator-activated receptor-gamma2 gene is associated with increased antilipolytic insulin sensitivity. *Diabetes,* 50, 876-81.

SUGAWARA, A., URUNO, A., KUDO, M., MATSUDA, K., YANG, C. W. & ITO, S. 2010. Effects of PPARgamma on hypertension, atherosclerosis, and chronic kidney disease. *Endocr J,* 57, 847-52.

TAN, N. S., MICHALIK, L., DESVERGNE, B. & WAHLI, W. 2005. Multiple expression control mechanisms of peroxisome proliferator-activated receptors and their target genes. *J Steroid Biochem Mol Biol,* 93, 99-105.

TANABE, K. K., LEMOINE, A., FINKELSTEIN, D. M., KAWASAKI, H., FUJII, T., CHUNG, R. T., LAUWERS, G. Y., KULU, Y., MUZIKANSKY, A., KURUPPU, D., LANUTI, M., GOODWIN, J. M., AZOULAY, D. & FUCHS, B. C. 2008. Epidermal growth factor gene functional polymorphism and the risk of hepatocellular carcinoma in patients with cirrhosis. *JAMA,* 299, 53-60.

TARRADE, A., SCHOONJANS, K., GUIBOURDENCHE, J., BIDART, J. M., VIDAUD, M., AUWERX, J., ROCHETTE-EGLY, C. & EVAIN-BRION, D. 2001a. PPAR gamma/RXR alpha heterodimers are involved in human CG beta synthesis and human trophoblast differentiation. *Endocrinology,* 142, 4504-14.

TARRADE, A., SCHOONJANS, K., PAVAN, L., AUWERX, J., ROCHETTE-EGLY, C., EVAIN-BRION, D. & FOURNIER, T. 2001b. PPARgamma/RXRalpha heterodimers control human trophoblast invasion. *J Clin Endocrinol Metab,* 86, 5017-24.

THADHANI, R., ECKER, J. L., MUTTER, W. P., WOLF, M., SMIRNAKIS, K. V., SUKHATME, V. P., LEVINE, R. J. & KARUMANCHI, S. A. 2004. Insulin resistance and alterations in angiogenesis: additive insults that may lead to preeclampsia. *Hypertension,* 43, 988-92.

THE GOPEC CONSORTIUM 2005. Disentangling fetal and maternal susceptibility for pre-eclampsia: a British multicenter candidate-gene study. *Am J Hum Genet,* 77, 127-31.

TROGSTAD, L., MAGNUS, P., SKJAERVEN, R. & STOLTENBERG, C. 2008. Previous abortions and risk of pre-eclampsia. *Int J Epidemiol,* 37, 1333-40.

TROGSTAD, L., SKRONDAL, A., STOLTENBERG, C., MAGNUS, P., NESHEIM, B. I. & ESKILD, A. 2004. Recurrence risk of preeclampsia in twin and singleton pregnancies. *Am J Med Genet A,* 126A, 41-5.

TSCHRITTER, O., STUMVOLL, M., MACHICAO, F., HOLZWARTH, M., WEISSER, M., MAERKER, E., TEIGELER, A., HARING, H. & FRITSCHE, A. 2002. The prevalent Glu23Lys polymorphism in the potassium inward rectifier 6.2 (KIR6.2) gene is associated with impaired glucagon suppression in response to hyperglycemia. *Diabetes,* 51, 2854-60.

VAN GAAL, L. F., GUTKIN, S. W. & NAUCK, M. A. 2008. Exploiting the antidiabetic properties of incretins to treat type 2 diabetes mellitus: glucagon-like peptide 1 receptor agonists or insulin for patients with inadequate glycemic control? *Eur J Endocrinol,* 158, 773-84.

VANWIJK, M. J., SVEDAS, E., BOER, K., NIEUWLAND, R., VANBAVEL, E. & KUBLICKIENE, K. R. 2002. Isolated microparticles, but not whole plasma, from women with preeclampsia impair endothelium-dependent relaxation in isolated myometrial arteries from healthy pregnant women. *Am J Obstet Gynecol,* 187, 1686-93.

VANWIJK, M. J., VANBAVEL, E., STURK, A. & NIEUWLAND, R. 2003. Microparticles in cardiovascular diseases. *Cardiovasc Res,* 59, 277-87.

VAULEON, E., AUGER, N., BENOUAICH-AMIEL, A., LAIGLE-DONADEY, F., KALOSHI, G., LEJEUNE, J., DELATTRE, J.-Y., THILLET, J. & SANSON, M. 2007. The 61 A/G EGF polymorphism is functional but is neither a prognostic marker nor a risk factor for glioblastoma. *Cancer Genetics and Cytogenetics,* 172, 33-37.

VENKATESHA, S., TOPORSIAN, M., LAM, C., HANAI, J., MAMMOTO, T., KIM, Y. M., BDOLAH, Y., LIM, K. H., YUAN, H. T., LIBERMANN, T. A., STILLMAN, I. E., ROBERTS, D., D'AMORE, P. A., EPSTEIN, F. H., SELLKE, F. W., ROMERO, R., SUKHATME, V. P., LETARTE, M. & KARUMANCHI, S. A. 2006. Soluble endoglin contributes to the pathogenesis of preeclampsia. *Nat Med,* 12, 642-9.

VERHAEGHE, J., VAN BREE, R., VAN HERCK, E., LAUREYS, J., BOUILLON, R. & VAN ASSCHE, F. A. 1993. C-peptide, insulin-like growth factors I and II, and insulin-like growth factor binding protein-1 in umbilical cord serum: correlations with birth weight. *Am J Obstet Gynecol,* 169, 89-97.

VERKAUSKIENE, R., FIGUERAS, F., DEGHMOUN, S., CHEVENNE, D., GARDOSI, J. & LEVY-MARCHAL, M. 2008. Birth weight and long-term metabolic outcomes: does the definition of smallness matter? *Horm Res,* 70, 309-15.

VISEL, A., ZHU, Y., MAY, D., AFZAL, V., GONG, E., ATTANASIO, C., BLOW, M. J., COHEN, J. C., RUBIN, E. M. & PENNACCHIO, L. A. 2010. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature,* 464, 409-12.

VON DADELSZEN, P., MAGEE, L. A. & ROBERTS, J. M. 2003. Subclassification of preeclampsia. *Hypertens Pregnancy,* 22, 143-8.

WAGNER, S. J., BARAC, S. & GAROVIC, V. D. 2007. Hypertensive pregnancy disorders: current concepts. *J Clin Hypertens (Greenwich),* 9, 560-6.

WAHLEN, K., SJOLIN, E. & HOFFSTEDT, J. 2008. The common rs9939609 gene variant of the fat mass- and obesity-associated gene FTO is related to fat cell lipolysis. *J Lipid Res,* 49, 607-11.

WAITE, L. L., LOUIE, R. E. & TAYLOR, R. N. 2005. Circulating activators of peroxisome proliferator-activated receptors are reduced in preeclamptic pregnancy. *J Clin Endocrinol Metab,* 90, 620-6.

WAITE, L. L., PERSON, E. C., ZHOU, Y., LIM, K. H., SCANLAN, T. S. & TAYLOR, R. N. 2000. Placental peroxisome proliferator-activated receptor-gamma is up-regulated by pregnancy serum. *J Clin Endocrinol Metab,* 85, 3808-14.

WANG, X., ZUCKERMAN, B., PEARSON, C., KAUFMAN, G., CHEN, C., WANG, G., NIU, T., WISE, P. H., BAUCHNER, H. & XU, X. 2002. Maternal Cigarette Smoking, Metabolic Gene Polymorphism, and Infant Birth Weight. *JAMA,* 287, 195-202.

WANG, X. L., OOSTERHOF, J. & DUARTE, N. 1999. Peroxisome proliferator-activated receptor gamma C161-->T polymorphism and coronary artery disease. *Cardiovasc Res,* 44, 588-94.

WANG, Y., TIAN, T., HU, Z., TANG, J., WANG, S., WANG, X., QIN, J., HUO, X., GAO, J., KE, Q., JIN, G., MA, H. & SHEN, H. 2008. EGF promoter SNPs, plasma EGF levels and risk of breast cancer in Chinese women. *Breast Cancer Res Treat,* 111, 321-7.

WATANABE, H. 1990. Epidermal growth factor in urine of pregnant women and in amniotic fluid throughout pregnancy. *Gynecol Endocrinol,* 4, 43-50.

WATANABE, R. M., ALLAYEE, H., XIANG, A. H., TRIGO, E., HARTIALA, J., LAWRENCE, J. M. & BUCHANAN, T. A. 2007. Transcription Factor 7-Like 2 (TCF7L2) Is Associated With Gestational Diabetes Mellitus and Interacts With Adiposity to Alter Insulin Secretion in Mexican Americans. *Diabetes,* 56, 1481-1485.

WEEDON, M. N., FRAYLING, T. M., SHIELDS, B., KNIGHT, B., TURNER, T., METCALF, B. S., VOSS, L., WILKIN, T. J., MCCARTHY, A., BEN-SHLOMO, Y., DAVEY SMITH, G., RING, S., JONES, R., GOLDING, J., TEAM, A. S., BYBERG, L., MANN, V., AXELSSON, T., SYVANEN, A.-C., LEON, D. & HATTERSLEY, A. T. 2005. Genetic Regulation of Birth Weight and Fasting Glucose by a Common Polymorphism in the Islet Cell Promoter of the Glucokinase Gene.

WEEDON, M. N., GLOYN, A. L., FRAYLING, T. M., HATTERSLEY, A. T., DAVEY SMITH, G. & BEN-SHLOMO, Y. 2003. Quantitative traits associated with the Type 2 diabetes susceptibility allele in Kir6.2. *Diabetologia,* 46, 1021-3.

WEI, J. N., SUNG, F. C., LI, C. Y., CHANG, C. H., LIN, R. S., LIN, C. C., CHIANG, C. C. & CHUANG, L. M. 2003. Low birth weight and high birth weight infants are both at an increased risk to have type 2 diabetes among schoolchildren in taiwan. *Diabetes Care,* 26, 343-8.

WELLS, A. 1999. EGF receptor. *Int J Biochem Cell Biol,* 31, 637-43.

WELLS, A., WELSH, J. B., LAZAR, C. S., WILEY, H. S., GILL, G. N. & ROSENFELD, M. G. 1990. Ligand-induced transformation by a noninternalizing epidermal growth factor receptor. *Science,* 247, 962-4.

WELTERS, H. J. & KULKARNI, R. N. 2008. Wnt signaling: relevance to beta-cell biology and diabetes. *Trends Endocrinol Metab,* 19, 349-55.

WHITLEY, G. S. & CARTWRIGHT, J. E. 2010. Cellular and molecular regulation of spiral artery remodelling: lessons from the cardiovascular field. *Placenta,* 31, 465-74.

WIEDEMANN, A., VOCKE, F., FITZGERALD, J. S., MARKERT, U. R., JESCHKE, U., LOHSE, P. & TOTH, B. 2010. Leptin gene (TTTC)(n) microsatellite polymorphism as well as leptin receptor R223Q and PPARgamma2 P12A substitutions are not associated with hypertensive disorders in pregnancy. *Am J Reprod Immunol,* 63, 310-7.

WIKSTROM, A. K., HAGLUND, B., OLOVSSON, M. & LINDEBERG, S. N. 2005. The risk of maternal ischaemic heart disease after gestational hypertensive disease. *BJOG,* 112, 1486-91.

WILLER, C. J., BONNYCASTLE, L. L., CONNEELY, K. N., DUREN, W. L., JACKSON, A. U., SCOTT, L. J., NARISU, N., CHINES, P. S., SKOL, A., STRINGHAM, H. M., PETRIE, J., ERDOS, M. R., SWIFT, A. J., ENLOE, S. T., SPRAU, A. G., SMITH, E., TONG, M., DOHENY, K. F., PUGH, E. W., WATANABE, R. M., BUCHANAN, T. A., VALLE, T. T., BERGMAN, R. N., TUOMILEHTO, J., MOHLKE, K. L., COLLINS, F. S. & BOEHNKE, M. 2007. Screening of 134 single nucleotide polymorphisms (SNPs) previously associated with type 2 diabetes replicates association with 12 SNPs in nine genes. *Diabetes,* 56, 256-64.

WILLIAMS, P. J., MISTRY, H. D., INNES, B. A., BULMER, J. N. & PIPKIN, F. B. 2010. Expression of AT1R, AT2R and AT4R and their roles in extravillous trophoblast invasion in the human. *Placenta,* 31, 448-55.

WINKVIST, A., MOGREN, I. & HOGBERG, U. 1998. Familial patterns in birth characteristics: impact on individual and population risks. *Int J Epidemiol,* 27, 248-54.

WOLF, M., HUBEL, C. A., LAM, C., SAMPSON, M., ECKER, J. L., NESS, R. B., RAJAKUMAR, A., DAFTARY, A., SHAKIR, A. S., SEELY, E. W., ROBERTS, J. M., SUKHATME, V. P., KARUMANCHI, S. A. & THADHANI, R. 2004. Preeclampsia and future cardiovascular disease: potential role of altered angiogenesis and insulin resistance. *J Clin Endocrinol Metab,* 89, 6239-43.

WRIGHT, J. K., DUNK, C. E., PERKINS, J. E., WINTERHAGER, E., KINGDOM, J. C. & LYE, S. J. 2006. EGF modulates trophoblast migration through regulation of Connexin 40. *Placenta,* 27 Suppl A, S114-21.

WTCCC. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature,* 447, 661-78.

WU, Y., ZHANG, Y. & ZHANG, J. 2005. Distribution of exonic splicing enhancer elements in human genes. *Genomics,* 86, 329-36.

YEN, C. J., BEAMER, B. A., NEGRI, C., SILVER, K., BROWN, K. A., YARNALL, D. P., BURNS, D. K., ROTH, J. & SHULDINER, A. R. 1997. Molecular scanning of the human peroxisome proliferator activated receptor gamma (hPPAR gamma) gene in diabetic Caucasians: identification of a Pro12Ala PPAR gamma 2 missense mutation. *Biochem Biophys Res Commun,* 241, 270-4.

YOUNG, B. C., LEVINE, R. J. & KARUMANCHI, S. A. 2010. Pathogenesis of preeclampsia. *Annu Rev Pathol,* 5, 173-92.

ZABENA, C., GONZALEZ-SANCHEZ, J. L., MARTINEZ-LARRAD, M. T., TORRES-GARCIA, A., ALVAREZ-FERNANDEZ-REPRESA, J., CORBATON-ANCHUELO, A., PEREZ-BARBA, M. & SERRANO-

RIOS, M. 2009. The FTO obesity gene. Genotyping and gene expression analysis in morbidly obese patients. *Obes Surg,* 19, 87-95.

ZATKOVA, A., MESSIAEN, L., VANDENBROUCKE, I., WIESER, R., FONATSCH, C., KRAINER, A. R. & WIMMER, K. 2004. Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. *Hum Mutat,* 24, 491-501.

ZEGGINI, E., PARKINSON, J. R., HALFORD, S., OWEN, K. R., WALKER, M., HITMAN, G. A., LEVY, J. C., SAMPSON, M. J., FRAYLING, T. M., HATTERSLEY, A. T. & MCCARTHY, M. I. 2005. Examining the relationships between the Pro12Ala variant in PPARG and Type 2 diabetes-related traits in UK samples. *Diabet Med,* 22, 1696-700.

ZEGGINI, E., WEEDON, M. N., LINDGREN, C. M., FRAYLING, T. M., ELLIOTT, K. S., LANGO, H., TIMPSON, N. J., PERRY, J. R., RAYNER, N. W., FREATHY, R. M., BARRETT, J. C., SHIELDS, B., MORRIS, A. P., ELLARD, S., GROVES, C. J., HARRIES, L. W., MARCHINI, J. L., OWEN, K. R., KNIGHT, B., CARDON, L. R., WALKER, M., HITMAN, G. A., MORRIS, A. D., DONEY, A. S., MCCARTHY, M. I. & HATTERSLEY, A. T. 2007. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science,* 316, 1336-41.

ZUKER, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.,* 31, 3406-3415.